



# STIC Search Report

## Biotech-Chem Library

STIC Database Tracking Number: 134960

TO: Maria Maryich  
Location: 2b84/2c70  
Art Unit: 1636  
October 22, 2004

Case Serial Number: 09/762648

From: P. Sheppard  
Location: Remsen Building  
Phone: (571) 272-2529

sheppard@uspto.gov

### Search Notes

BEST AVAILABLE COPY

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 19, 2004, 13:19:25 ; Search time 1557 Seconds  
(without alignments)  
303.723 Million cell updates/sec

Title: US-09-762-648-4  
Perfect score: 10  
Sequence: 1 UGCDGSHMD 10

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 2364489745 residues  
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmb1:\*  
1: gb\_ha:\*  
2: gb\_hg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sbs:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6.2	62.0	10	6	ARI07856 Sequence
2	6.2	62.0	10	6	BD238641 Sequence
3	6.2	62.0	10	6	BD238641 Preparati
4	6.2	62.0	10	6	CQ766701 Sequence
5	6.2	62.0	10	6	CQ828869 Sequence
6	6.2	62.0	10	6	AR234537 Sequence
7	6.2	62.0	10	6	AR303297 Sequence
8	6.2	62.0	11	6	AI3081 Nucleotide
9	6.2	62.0	11	6	AR074496 Sequence
10	6.2	62.0	11	6	AR081176 Sequence
11	6.2	62.0	11	6	AR085373 Sequence
12	6.2	62.0	11	6	AR088121 Sequence
13	6.2	62.0	11	6	ARI04280 Sequence
14	6.2	62.0	11	6	ARI39988 Sequence
15	6.2	62.0	11	6	ARI40307 Sequence
16	6.2	62.0	11	6	ARI40585 Sequence
17	6.2	62.0	11	6	ARI43542 Sequence
18	6.2	62.0	11	6	ARI71448 Sequence
19	6.2	62.0	11	6	ARI71619 Sequence

20	6.2	62.0	11	6	BD243209 MN gene a
21	6.2	62.0	11	6	CQ828868 Sequence
22	6.2	62.0	11	6	CQ833562 Sequence
23	6.2	62.0	11	6	CQ836876 Sequence
24	6.2	62.0	11	6	CQ837108 Sequence
25	6.2	62.0	11	6	CQ837113 Sequence
26	6.2	62.0	11	6	CQ837507 Sequence
27	6.2	62.0	11	6	AR301722 Sequence
28	6.2	62.0	11	6	AX225318 Sequence
29	6.2	62.0	11	6	AX339223 Sequence
30	6.2	62.0	11	6	AX393232 Sequence
31	6.2	62.0	11	6	AX470780 Sequence
32	6.2	62.0	11	6	AX470844 Sequence
33	6.2	62.0	11	6	AX471086 Sequence
34	6.2	62.0	11	6	AX624499 Sequence
35	6.2	62.0	11	6	AX627640 Sequence
36	6.2	62.0	11	6	AX627775 Sequence
37	6.2	62.0	11	6	AX627871 Sequence
38	6.2	62.0	11	6	AX628442 Sequence
39	6.2	62.0	11	6	AX629146 Sequence
40	6.2	62.0	11	6	AX629397 Sequence
41	6.2	62.0	11	6	AX630172 Sequence
42	6.2	62.0	11	6	AX630221 Sequence
43	6.2	62.0	11	6	AX630249 Sequence
44	6.2	62.0	11	6	AX631920 Sequence
45	6.2	62.0	11	6	AX632797 Sequence

## ALIGNMENTS

RESULT 1	ARI07856	Sequence 102 from patent US 6110667.	10 bp	DNA	linear	PAT 14-FEB-2001
LOCUS	ARI07856					
DEFINITION	Sequence 102 from patent US 6110667.					
ACCESSION	ARI07856					
VERSION	ARI07856.1	GI:12823343				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	1 (bases 1 to 10)					
AUTHORS	Lopez-Nieto C, Eduardo, and Nigam S, Kumar.					
TITLE	Processes, apparatus and compositions for characterizing nucleotide sequences based on K-tuple analysis					
JOURNAL	Patent: US 6110667-A 102 29-AUG-2000;					
FEATURES	Location/Qualifiers					
source	1..10					
ORIGIN	/organism="unknown"					
	/mol_type="unassigned DNA"					
Query Match	Best Local Similarity 40.0%; Pred. No. 1.4e+07;					
Matches	4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;					
Qy	1 UGCDGSHMD 10					
Db	1 TGCTGTCAT 10					
RESULT 2	BD238641/c					
LOCUS	BD238641					
DEFINITION	Preparation and use of superior vaccines.					
ACCESSION	BD238641					
VERSION	BD238641.1	GI:33048411				
KEYWORDS	UP 2002534056-A/59.					
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE	1 (bases 1 to 10)					
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					

AUTHORS Roberte, B.L. and Shankara, S.  
TITLE Preparation and use of superior vaccines  
JOURNAL Patent: JP 2002534056-A 59 15-OCT-2002;  
GENZYME CORP  
COMMENT OS Homo sapiens (human)  
PN JP 2002534056-A/59  
PD 15-OCT-2002  
PR 19-JUN-1998 US 60/090039, 19-JUN-1998 US 60/090040 PR  
19-JUN-1998 US 60/090041, 19-JUN-1998 US 60/089853 PR  
19-JUN-1998 US 60/089997, 19-JUN-1998 US 60/090079 PR  
19-JUN-1998 US 60/090035, 19-JUN-1998 US 60/089993 PR  
19-JUN-1998 US 60/089992, 19-JUN-1998 US 60/090072 PR  
19-JUN-1998 US 60/089878, 19-JUN-1998 US 60/089991 PR  
19-JUN-1998 US 60/090000, 19-JUN-1998 US 60/090048 PR  
19-JUN-1998 US 60/090042, 19-JUN-1998 US 60/090036 PR  
19-JUN-1998 US 60/090044, 19-JUN-1998 US 60/090036 PR  
19-JUN-1998 US 60/090042, 19-JUN-1998 US 60/090036 PR  
19-JUN-1998 US 60/090044, 19-JUN-1998 US 60/089844 PR  
19-JUN-1998 US 60/090080, 19-JUN-1998 US 60/089833 PR  
19-JUN-1998 US 60/089994, 19-JUN-1998 US 60/090077 PR  
19-JUN-1998 US 60/090078, 19-JUN-1998 US 60/090047 PR  
08-DEC-1998 US 60/090076, 19-JUN-1998 US 60/090045 PR  
08-DEC-1998 US 60/111715  
PI BRUCE L. ROBERTS, SRINIVAS SHANKARA  
PC C12N1/19, C12N5/09, A61K39/00, A61P35/00, A61P37/04, C12N1/15, PC  
C12N1/21, C12N5/10, G01N33/15, G01N33/50, G01N33/53, G01N33/566, PC  
G01N37/00  
PC C12N15/00, C12N5/00, C12N15/00  
CC Preparation and use of superior vaccines  
FH Key Location/Qualifiers  
FT source 1.10  
Location/Qualifiers  
1.10  
/organism="Homo sapiens (human)".  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 62.0%; Score 6.2; DB 6; Length 10;  
Best Local Similarity 40.0%; Pred. No. 1.4e+07;  
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 UGCDDGHNMD 10  
Db 10 TGCCTGGCTAG 1

RESULT 3  
BD238894/c 10 bp DNA linear PAT 17-JUL-2003  
LOCUS Preparation and use of superior vaccines.  
DEFINITION BD238894  
ACCESSION BD238894  
VERSION BD238894.1 GI:33048664  
KEYWORDS UP 2002534056-A/312.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 10)  
Robert, B.L. and Shankara, S.  
AUTHORS Preparation and use of superior vaccines  
TITLE Patent: JP 2002534056-A 312 15-OCT-2002;  
JOURNAL GENZYME CORP  
COMMENT OS Homo sapiens (human)  
PN JP 2002534056-A/312  
PD 15-OCT-2002  
PR 18-JUN-1998 US 60/090039, 19-JUN-1998 US 60/090040 PR  
19-JUN-1998 US 60/090041, 19-JUN-1998 US 60/089853 PR  
19-JUN-1998 US 60/089997, 19-JUN-1998 US 60/090079 PR  
19-JUN-1998 US 60/090035, 19-JUN-1998 US 60/089993 PR

19-JUN-1998 US 60/089992, 19-JUN-1998 US 60/090072 PR  
19-JUN-1998 US 60/089878, 19-JUN-1998 US 60/089991 PR  
19-JUN-1998 US 60/090000, 19-JUN-1998 US 60/090048 PR  
19-JUN-1998 US 60/089999, 19-JUN-1998 US 60/090043 PR  
19-JUN-1998 US 60/090042, 19-JUN-1998 US 60/090036 PR  
19-JUN-1998 US 60/090044, 19-JUN-1998 US 60/090036 PR  
19-JUN-1998 US 60/090042, 19-JUN-1998 US 60/090036 PR  
19-JUN-1998 US 60/090044, 19-JUN-1998 US 60/089844 PR  
19-JUN-1998 US 60/090080, 19-JUN-1998 US 60/089833 PR  
19-JUN-1998 US 60/089994, 19-JUN-1998 US 60/090077 PR  
19-JUN-1998 US 60/090078, 19-JUN-1998 US 60/090047 PR  
08-DEC-1998 US 60/111715  
PI BRUCE L. ROBERTS, SRINIVAS SHANKARA  
PC C12N15/09, C12N15/09, A61K39/00, A61P35/00, A61P37/04, C12N1/15, PC  
C12N1/21, C12N5/10, G01N33/15, G01N33/50, G01N33/53, G01N33/566, PC  
G01N37/00,  
PC C12N15/00, C12N5/00, C12N15/00  
CC Preparation and use of superior vaccines  
FH Key Location/Qualifiers  
FT source 1.10  
Location/Qualifiers  
1.10  
/organism="Homo sapiens (human)".  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 62.0%; Score 6.2; DB 6; Length 10;  
Best Local Similarity 40.0%; Pred. No. 1.4e+07;  
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 UGCDDGHNMD 10  
Db 10 TGCAGAGAAA 1

RESULT 4  
CQ766701 10 bp DNA linear PAT 03-MAR-2004  
LOCUS Sequence 57 from Patent WO2004005541.  
DEFINITION CQ766701  
ACCESSION CQ766701  
VERSION CQ766701.1 GI:44908931  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
artificial sequences.  
REFERENCE 1  
AUTHORS van Broeckhoven, C., de Jonghe, P., Timmerman, V. and Verhoeven, K.  
TITLE Diagnostic tests for the detection of peripheral neuropathy  
JOURNAL Vlaams Interuniversitair Instituut voor Biotechnologie vzw, w. (BE)  
FEATURES  
SOURCE 1.10  
Location/Qualifiers  
1.10  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="3-intron/exon, exon 2, gene RAB7"

ORIGIN  
Query Match 62.0%; Score 6.2; DB 6; Length 10;  
Best Local Similarity 40.0%; Pred. No. 1.4e+07;  
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 UGCDDGHNMD 10  
Db 1 TGCAGGTAAAG 10

RESULT 5  
CQ828869/c 10 bp DNA linear PAT 05-JUL-2004  
LOCUS Sequence 587 from Patent WO2004053120.  
DEFINITION

```

ACCESSION      CQ828865
VERSION        CQ828865.1
KEYWORDS
SOURCE
ORGANISM      Homo sapiens (human)
              Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS      1
              Weihe,E., Bieller,A. and Schaefer,M.K.
              Regulatory elements in the 5' region of the vrl gene
              Patent: WO 2004053120-A 587 24-JUN-2004;
              Gruenenthal GmbH (DE)
FEATURES
source
              1..10
              /organism="Homo sapiens"
              /mol_type="unassigned DNA"
              /db_xref="taxon:9606"
              /note="V$MYOD Q6"
ORIGIN
Query Match      62.0%; Score 6.2; DB 6; Length 10;
Best Local Similarity 40.0%; Pred.No. 1.4e+07;
Matches      4; Conservative      5; Mismatches      1; Indels      0; Gaps      0;
QY      1 UGCDGSHNMD 10
              :::::
              10 TGCAGGTGAA 1
Db
RESULT 6
AR234537
LOCUS      AR234537      10 bp      DNA      PAT 20-DEC-2002
DEFINITION      Sequence 71 from patent US 6458584.
ACCESSION      AR234537
VERSION      AR234537.1 GI:272727241
KEYWORDS
SOURCE      Unknown.
ORGANISM      Unknown.
REFERENCE      1 (bases 1 to 10)
              Mirzabekov,A., Guschin,D.Y., Chik,V., Drobyshev,A., Fotin,A.,
              Yershov,G. and Lysov,Y.
              Customized oligonucleotide microchips that convert multiple genetic
              information to simple patterns, are portable and reusable
              Patent: US 6458584-A 71 01-OCT-2002;
              Location/Qualifiers
              1..10
              /organism="unknown"
              /mol_type="genomic DNA"
ORIGIN
Query Match      62.0%; Score 6.2; DB 6; Length 10;
Best Local Similarity 40.0%; Pred.No. 1.4e+07;
Matches      4; Conservative      5; Mismatches      1; Indels      0; Gaps      0;
QY      1 UGCDGSHNMD 10
              :::::
              1 TGCAGGTGAA 10
Db
RESULT 7
AR303297
LOCUS      AR303297      10 bp      DNA      PAT 12-JUN-2003
DEFINITION      Sequence 22 from patent US 6544736.
ACCESSION      AR303297
VERSION      AR303297.1 GI:31692073
KEYWORDS
SOURCE      Unknown.
ORGANISM      Unknown.
REFERENCE      1 (bases 1 to 10)
              Shimamoto,A., Furuichi,Y., Shibata,Y., Funaki,H., Ohara,E. and
              Watahiki,M.

```

```

TITLE          Method for synthesizing cDNA from mRNA sample
JOURNAL        Patent: US 6544736-A 22 08-APR-2003;
FEATURES
  source
    1..10
    /organism="unknown"
    /mol_type="genomic DNA"
ORIGIN
Query Match      62.0%; Score 6.2; DB 6; Length 10;
Best Local Similarity 40.0%; Pred. No. 1.4e+07;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
QY
  1 UGCDGSHNMD 10
  :||:|:|:|:
  Db          1 TGCTGGAAAG 10
RESULT 8
A13081
LOCUS          A13081 11 bp DNA linear PAT 08-FEB-1994
DEFINITION    Nucleotide sequence 36 from patent number EP0353188.
ACCESSION     A13081
VERSION       A13081.1 GI:489599
KEYWORDS
SOURCE        unidentified
              unidentified.
ORGANISM      unidentified.
REFERENCE     1 (bases 1 to 11)
AUTHORS      Heim,J., Meyhack,B. and Visser,J.
TITLE        Novel expression system
JOURNAL      Patent: EP 0353188-A 36 31-JAN-1990;
              CIBA-GEIGY AG
FEATURES
  source
    1..11
    /organism="unidentified"
    /mol_type="unassigned DNA"
    /db_xref="taxon:32644"
ORIGIN
Query Match      62.0%; Score 6.2; DB 6; Length 11;
Best Local Similarity 40.0%; Pred. No. 1.3e+07;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
QY
  1 UGCDGSHNMD 10
  :||:|:|:|:
  Db          2 TGCAGGTCAA 11
RESULT 9
A074496
LOCUS          A074496 11 bp DNA / linear PAT 28-AUG-2000
DEFINITION    Sequence 75 from patent US 5955075.
ACCESSION     A074496
VERSION       A074496.1 GI:10001251
KEYWORDS
SOURCE        Unknown.
              Unknown.
ORGANISM      Unclassified.
REFERENCE     1 (bases 1 to 11)
AUTHORS      Zavada,J., Pastorekova,S. and Pastorek,J.
TITLE        Method of inhibiting tumor growth using antibodies to MN protein
JOURNAL      Patent: US 5955075-A 75 21-SEP-1999;
              Location/Qualifiers
FEATURES
  source
    1..11
    /organism="unknown"
    /mol_type="unassigned DNA"
ORIGIN
Query Match      62.0%; Score 6.2; DB 6; Length 11;
Best Local Similarity 40.0%; Pred. No. 1.3e+07;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
QY
  1 UGCDGSHNMD 10

```

	Db	1 TGCTGTGATAG 10 : : : : : : :			
	RESULT 10				
	LOCUS	AR081176	11 bp	DNA	linear
	DEFINITION	Sequence 75 from patent US 5972353.			PAT 31-AUG-2000
	ACCESSION	AR081176			
	VERSION	AR081176.1 GI:10007904			
	KEYWORDS	.			
	SOURCE	Unknown.			
	ORGANISM	Unclassified.			
	REFERENCE	1 (bases 1 to 11)			
	AUTHORS	Zavada,J., Pastorekova,S. and Pastorek,J.			
	TITLE	MN protein, polypeptides, fusion proteins and fusion polypeptides			
	JOURNAL	Patent: US 5972353-A 75 26-OCT-1999;			
	FEATURES	Location/Qualifiers			
	source	1..11			
		/organism="unknown"			
	ORIGIN	/mol_type="unassigned DNA"			
	Query Match	62.0%; Score 6.2; DB 6; Length 11;			
	Best Local Similarity	40.0%; Pred.No. 1.3e+07;			
	Matches	4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;			
	Oy	1 UGCDCGHNMID 10 :    ::: :: :			
	Db	1 TGCTGTGATAG 10			
	RESULT 11				
	LOCUS	AR085373	11 bp	DNA	linear
	DEFINITION	Sequence 75 from patent US 5981711.			PAT 01-SEP-2000
	ACCESSION	AR085373			
	VERSION	AR085373.1 GI:10012142			
	KEYWORDS	.			
	SOURCE	Unknown.			
	ORGANISM	Unclassified.			
	REFERENCE	1 (bases 1 to 11)			
	AUTHORS	Zavada,J., Pastorekova,S. and Pastorek,J.			
	TITLE	MN-specific antibodies and hybridomas			
	JOURNAL	Patent: US 5981711-A 75 09-NOV-1999;			
	FEATURES	Location/Qualifiers			
	source	1..11			
		/organism="unknown"			
	ORIGIN	/mol_type="unassigned DNA"			
	Query Match	62.0%; Score 6.2; DB 6; Length 11;			
	Best Local Similarity	40.0%; Pred.No. 1.3e+07;			
	Matches	4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;			
	Oy	1 UGCDCGHNMID 10 :    ::: :: :			
	Db	1 TGCTGTGATAG 10			
	RESULT 12				
	LOCUS	AR088121	11 bp	DNA	linear
	DEFINITION	Sequence 75 from patent US 5989838.			PAT 07-SEP-2000
	ACCESSION	AR088121			
	VERSION	AR088121.1 GI:10014884			
	KEYWORDS	Unknown.			
	SOURCE	Unknown.			
	ORGANISM	Unclassified.			
	REFERENCE	1 (bases 1 to 11)			

```

AUTHORS      Zavada,J., Pastorekova,S. and Pastorek,J.
TITLE        Immunological methods of detecting MN proteins and MN polypeptides
JOURNAL      Patent: US 5969838-A 75 23-NOV-1999;
FEATURES     Location/Qualifiers
            source
            1..11
            /organism="unknown"
            /mol_type="unassigned DNA"
ORIGIN
Query Match      62.0%; Score 6.2; DB 6; Length 11;
Best Local Similarity 40.0%; Pred. No. 1.3e+07;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
QY 1 UGCDGGHND 10
      :|::|::|:
      1 TGCTGTGTAG 10

RESULT 13
LOCUS      AR104280                11 bp      DNA      PAT 14-FEB-2001
DEFINITION Sequence 75 from patent US 6093548.
ACCESSION  AR104280
VERSION    AR104280.1 GI:12816988
KEYWORDS
SOURCE     Unknown.
ORGANISM   Unknown.
REFERENCE   Unclassified.
            1 (bases 1 to 11)
            Zavada,J., Pastorekova,S. and Pastorek,J.
            Detection and quantitation of MN-specific antibodies
            Patent: US 6093548-A 75 25-JUL-2000;
            Location/Qualifiers
            1..11
            /organism="unknown"
            /mol_type="unassigned DNA"
ORIGIN
Query Match      62.0%; Score 6.2; DB 6; Length 11;
Best Local Similarity 40.0%; Pred. No. 1.3e+07;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
QY 1 UGCDGGHND 10
      :|::|::|:
      1 TGCTGTGTAG 10

Db 1 TGCTGTGTAG 10

RESULT 14
LOCUS      AR139986/c              11 bp      DNA      PAT 16-JUN-2001
DEFINITION Sequence 99 from patent US 6207417.
ACCESSION  AR139986
VERSION    AR139986.1 GI:14482484
KEYWORDS
SOURCE     Unknown.
ORGANISM   Unknown.
REFERENCE   Unclassified.
            1 (bases 1 to 11)
            Zebro,K.M., Bosselman,R.A., Suggs,S.V. and Martin,F.H.
            DNA encoding stem cell factor
            Patent: US 6207417-A 99 27-MAR-2001;
            Location/Qualifiers
            1..11
            /organism="unknown"
            /mol_type="unassigned DNA"
ORIGIN
Query Match      62.0%; Score 6.2; DB 6; Length 11;
Best Local Similarity 40.0%; Pred. No. 1.3e+07;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
QY 1 UGCDGGHND 10
      :|::|::|:
      1 TGCTGTGTAG 10

```

Db 11 TGCAGAGAT 2

RESULT 15  
ARI40307/c

LOCUS ARI40307 11 bp DNA linear PAT 16-JUN-2001  
DEFINITION Sequence 99 from patent US 6207454.

ACCESSION ARI40307  
VERSION ARI40307.1 GI:14482803

KEYWORDS  
SOURCE Unknown.

ORGANISM Unknown.  
Unclassified.

REFERENCE 1 (bases 1 to 11)  
AUTHORS Zeebo,K.M., Bosseiman,R.A., Suggs,S.V. and Martin,F.H.  
TITLE Method for enhancing the efficiency of gene transfer with stem cell  
factor (SCF) polypeptide

JOURNAL Patent: US 6207454-A 99 27-MAR-2001;  
FEATURES Location/Qualifiers

1..11  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Query Match 62.0%; Score 6.2; DB 6; Length 11;  
Best Local Similarity 40.0%; Pred.No.1.3e+07;  
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 UGCDGHNMD 10  
:|:|:|:|:  
Db 11 TGCAGAGAT 2

Search completed: October 19, 2004, 15:31:01  
Job time : 1560 secs

BEST AVAILABLE COPY

This Page Blank (uspto)



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 19, 2004, 11:15:35 : Search time 230 Seconds  
(without alignments)  
181.015 Million cell updates/sec

Title: US-09-762-648-4  
Perfect score: 10  
Sequence: 1 UGCDGQHMD 10

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues  
Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_23Sep04:\*

- 1: geneeqn1980s:\*
- 2: geneeqn1990s:\*
- 3: geneeqn2000s:\*
- 4: geneeqn2001as:\*
- 5: geneeqn2001bs:\*
- 6: geneeqn2002as:\*
- 7: geneeqn2002bs:\*
- 8: geneeqn2003as:\*
- 9: geneeqn2003bs:\*
- 10: geneeqn2003cs:\*
- 11: geneeqn2003ds:\*
- 12: geneeqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	6.2	62.0	10	2	AAQ96908 HIV-1 NL4
C 2	6.2	62.0	10	2	AAQ96474 HIV-1 NL4
C 3	6.2	62.0	10	2	AAQ96474 HIV-1 NL4
C 4	6.2	62.0	10	2	AAQ96474 HIV-1 NL4
C 5	6.2	62.0	10	2	AAQ96474 HIV-1 NL4
C 6	6.2	62.0	10	2	AAQ96474 HIV-1 NL4
C 7	6.2	62.0	10	2	AAQ96474 HIV-1 NL4
C 8	6.2	62.0	10	2	AAQ96474 HIV-1 NL4
C 9	6.2	62.0	10	2	AAQ96474 HIV-1 NL4
C 10	6.2	62.0	10	2	AAQ96474 HIV-1 NL4
C 11	6.2	62.0	10	2	AAQ96474 HIV-1 NL4
C 12	6.2	62.0	10	2	AAQ96474 HIV-1 NL4
C 13	6.2	62.0	10	2	AAQ96474 HIV-1 NL4
C 14	6.2	62.0	10	2	AAQ96474 HIV-1 NL4
C 15	6.2	62.0	10	2	AAQ96474 HIV-1 NL4
C 16	6.2	62.0	10	2	AAQ96474 HIV-1 NL4
C 17	6.2	62.0	10	2	AAQ96474 HIV-1 NL4
C 18	6.2	62.0	10	2	AAQ96474 HIV-1 NL4
C 19	6.2	62.0	10	2	AAQ96474 HIV-1 NL4
C 20	6.2	62.0	10	2	AAQ96474 HIV-1 NL4
C 21	6.2	62.0	10	2	AAQ96474 HIV-1 NL4

C 22	6.2	62.0	10	6	ABK92648 Primer-ex
C 23	6.2	62.0	10	6	ABL52050 Human SLC
C 24	6.2	62.0	10	6	AAI6764 Human APO
C 25	6.2	62.0	10	6	ABK96151 Human CYP
C 26	6.2	62.0	10	6	ADH22211 Primer ex
C 27	6.2	62.0	10	10	ADD71266 Mouse ET
C 28	6.2	62.0	10	10	ADD71292 Human ET
C 29	6.2	62.0	10	10	ACA63200 Human AD
C 30	6.2	62.0	10	10	ABX15890 Human BUB
C 31	6.2	62.0	10	12	ADN89079 Human VRI
C 32	6.2	62.0	10	12	ADQ30294 Human VRI
C 33	6.2	62.0	11	2	AAI18993 Murine MR
C 34	6.2	62.0	11	3	AAI13719 Rat stem
C 35	6.2	62.0	11	3	AAI16597 Human MN
C 36	6.2	62.0	11	3	AAI2516 Human MN
C 37	6.2	62.0	11	3	AAI92135 Human Lnx
C 38	6.2	62.0	11	4	AAH41359 Recombina
C 39	6.2	62.0	11	4	AAH41359 Recombina
C 40	6.2	62.0	11	4	AAH41359 Recombina
C 41	6.2	62.0	11	4	AAH41359 Recombina
C 42	6.2	62.0	11	5	AAH23917 Stem cell
C 43	6.2	62.0	11	5	AAH23917 Stem cell
C 44	6.2	62.0	11	5	AAH23917 Stem cell
C 45	6.2	62.0	11	6	ABQ86666 Human skl

## ALIGNMENTS

RESULT 1	AAQ96908/C	AAQ96908 standard; DNA; 10 BP.
XX	AAQ96908;	
AC	AAQ96908;	
XX	AAQ96908;	
DT	16-OCT-2003 (revised)	
DT	26-MAR-1996 (first entry)	
XX	HIV-1 NL4-3 nef gene nucleotide deletion 503.	
XX	HIV-1; AIDS; attenuation; vaccine; nef gene; avirulence; ss.	
OS	Human immunodeficiency virus 1.	
XX	W09521912-AL.	
PD	17-AUG-1995.	
XX	14-FEB-1995;	95MO-AU000063.
XX	14-FEB-1994;	94AU-00003864.
PR	21-FEB-1994;	94AU-00004002.
PR	23-DEC-1994;	94AU-00000284.
PA	(MACF-) MACFARLANE BURNET CENT MEDICAL.	
PA	(AURE-) AUSTRALIAN RED CROSS SOC.	
XX	Deacon NJ, Learmont JC, Mcphee DA, Crowe S, Cooper D;	
XX	WPI, 1995-293115/38.	
DR	New non-pathogenic HIV-1 strain carrying a deletion in its nef gene or	
PT	LTR region - can be used in a vaccine to inhibit/reduce productive	
PT	infection in an individual by a pathogenic strain.	
XX	Claim 13; Page 194; 301pp; English.	
XX	Attenuation of pathogenic HIV-1 strain NL4-3 involves deletion of 1 or	
CC	more decaunucleotides (AAQ96406-Q97018) from the nef gene and/or 1 or more	
CC	decaunucleotides (AAQ97019-Q97166) from the LTR region; the sequence of	
CC	AAQ96406 corresponds to nucleotides 1-10 of the nef gene (AAQ96141). The	
CC	resulting avirulent HIV strains are still capable of inducing an immune	
CC	response in humans, and enable the generation of therapeutic, diagnostic	

CC and targeting agents against HIV-1 infection. (Updated on 16-OCT-2003 to  
 CC standardise OS field)  
 XX  
 SQ Sequence 10 BP; 2 A; 3 C; 3 G; 2 T; 0 U; 0 Other;

Query Match 62.0%; Score 6.2; DB 2; Length 10;  
 Best Local Similarity 40.0%; Pred. No. 6.7e+05;  
 Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 UCCDGGHND 10  
 :|||: :  
 Db 10 TGCAGCTCA 1

## RESULT 2

AA096474/C  
 ID AA096474 standard; DNA; 10 BP.

XX AA096474;

AC 16-OCT-2003 (revised)

DT 20-MAR-1996 (first entry)

XX HIV-1 NL4-3 nef gene nucleotide deletion 69.

XX HIV-1; AIDS; attenuation; vaccine; nef gene; avirulence; ss.

XX Human immunodeficiency virus 1.

XX WO9521912-A1.

XX 17-AUG-1995.

XX 14-FEB-1995; 95WO-AU000063.

XX 14-FEB-1994; 94AU-00003864.

XX 21-FEB-1994; 94AU-00004002.

XX 23-DEC-1994; 94AU-00000284.

XX (MACF-) MACFARLANE BURNET CENT MEDICAL.

XX (AURE-) AUSTRALIAN RED CROSS SOC.

XX Deacon NJ, Learmont JC, Mcphee DA, Crowe S, Cooper D;

XX WPI; 1995-293115/38.

XX New non-pathogenic HIV-1 strain carrying a deletion in its nef gene or  
 PT LTR region - can be used in a vaccine to inhibit/reduce productive  
 PT infection in an individual by a pathogenic strain.

XX Claim 13; Page 189; 301pp; English.

XX Attenuation of pathogenic HIV-1 strain NL4-3 involves deletion of 1 or  
 CC more deannucleotides (AA096406-Q97018) from the nef gene and/or 1 or more  
 CC deannucleotides (AA097019-Q97166) from the LTR region; the sequence of  
 CC AA096406 corresponds to nucleotides 1-10 of the nef gene (AA096141). The  
 CC resulting avirulent HIV strains are still capable of inducing an immune  
 CC response in humans, and enable the generation of therapeutic, diagnostic  
 CC and targeting agents against HIV-1 infection. (Updated on 16-OCT-2003 to  
 CC standardise OS field)

XX Sequence 10 BP; 3 A; 3 C; 3 G; 1 T; 0 U; 0 Other;

XX Query Match 62.0%; Score 6.2; DB 2; Length 10;  
 Best Local Similarity 40.0%; Pred. No. 6.7e+05;  
 Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 UCCDGGHND 10  
 :|||: :  
 Db 10 TGCAGCTCA 1

## RESULT 3

AA096474/C  
 ID AA096474 standard; DNA; 10 BP.

XX AA096474;

AC 16-OCT-2003 (revised)

DT 20-MAR-1996 (first entry)

XX HIV-1 NL4-3 nef gene nucleotide deletion 69.

XX HIV-1; AIDS; attenuation; vaccine; nef gene; avirulence; ss.

XX Human immunodeficiency virus 1.

XX WO9521912-A1.

XX 17-AUG-1995.

XX 14-FEB-1995; 95WO-AU000063.

XX 14-FEB-1994; 94AU-00003864.

XX 21-FEB-1994; 94AU-00004002.

XX 23-DEC-1994; 94AU-00000284.

XX (MACF-) MACFARLANE BURNET CENT MEDICAL.

XX (AURE-) AUSTRALIAN RED CROSS SOC.

XX Deacon NJ, Learmont JC, Mcphee DA, Crowe S, Cooper D;

XX WPI; 1995-293115/38.

XX New non-pathogenic HIV-1 strain carrying a deletion in its nef gene or  
 PT LTR region - can be used in a vaccine to inhibit/reduce productive  
 PT infection in an individual by a pathogenic strain.

XX Claim 13; Page 189; 301pp; English.

XX Attenuation of pathogenic HIV-1 strain NL4-3 involves deletion of 1 or  
 CC more deannucleotides (AA096406-Q97018) from the nef gene and/or 1 or more  
 CC deannucleotides (AA097019-Q97166) from the LTR region; the sequence of  
 CC AA096406 corresponds to nucleotides 1-10 of the nef gene (AA096141). The  
 CC resulting avirulent HIV strains are still capable of inducing an immune  
 CC response in humans, and enable the generation of therapeutic, diagnostic  
 CC and targeting agents against HIV-1 infection. (Updated on 16-OCT-2003 to  
 CC standardise OS field)

XX Sequence 10 BP; 3 A; 3 C; 3 G; 1 T; 0 U; 0 Other;

XX Query Match 62.0%; Score 6.2; DB 2; Length 10;  
 Best Local Similarity 40.0%; Pred. No. 6.7e+05;  
 Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 UCCDGGHND 10  
 :|||: :  
 Db 10 TGCAGCTCA 1

AAT29372  
 ID AAT29372 standard; DNA; 10 BP.

XX AAT29372;

AC 25-MAR-2003 (revised)

DT 28-JUN-1996 (first entry)

XX 5'-primer for mammalian G-protein coupled receptor coding sequences.

XX 5'-primer; mammalian; G-protein coupled receptor; PCR primer kit;

XX characterisation; biological samples; PCR amplification; indexing;

XX identification; cloning; analysis; genes; genome mapping;

XX disease diagnosis; ss.

XX Synthetic.

XX WO9531574-A1.

XX 23-NOV-1995.

XX 12-MAY-1995; 95WO-US006032.

XX 16-MAY-1994; 94US-00242887.

XX (BGHM ) BRIGHAM & WOMENS HOSPITAL.

XX Lopeznieto CE, Nigam SK;

XX WPI; 1996-010958/01.

XX Characterisation of nucleotide sequences using primer pairs - by PCR  
 PT amplification and indexing of amplification prods. w.r.t. primers used  
 PT for genome mapping and disease diagnosis.

XX Claim 46; Page 55; 72pp; English.

XX The 5'-primers AAT29262-382, and the complementary 3'-primers derived  
 CC from them, which target mammalian G-protein coupled receptor coding  
 CC sequences, together comprise a PCR primer kit. The kit is used in a new  
 CC method for the characterisation of nucleic acid sequences obtd. from  
 CC mammalian biological samples, which comprises PCR amplification and  
 CC indexing of the prods. w.r.t. the primer pair that hybridised to its  
 CC delineating subsequences. The method may be used in the identification,  
 CC cloning and analysis of genes, e.g. in genome mapping, and disease  
 CC diagnosis. (Updated on 25-MAR-2003 to correct PI field.)

XX Sequence 10 BP; 1 A; 2 C; 3 G; 4 T; 0 U; 0 Other;

XX Query Match 62.0%; Score 6.2; DB 2; Length 10;  
 Best Local Similarity 40.0%; Pred. No. 6.7e+05;  
 Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 UCCDGGHND 10  
 :|||: :  
 Db 1 TCGTGGTCAT 10

## RESULT 4

AAV45971  
 ID AAV45971 standard; DNA; 10 BP.

XX AAV45971;

AC 08-OCT-1998 (first entry)

DT Biosensor oligonucleotide #71.

XX Biosensor; microchip; detection; identification; hybridisation; ss.

XX Synthetic.

XX WO9828444-A2.

```

XX 02-JUL-1998.
PD
XX
XX 19-DEC-1997; 97WO-US023778.
PF
XX
XX 23-DEC-1996; 96US-00780026.
PR
XX
XX (UYCH-) UNIV CHICAGO.
PA
XX
XX Mitzabekov A, Guschin DY, Shik V, Potin A, Yershov G, Lysov Y;
PI
XX WPI; 1998-377672/32.
DR
XX
XX New customised oligonucleotide microchips - comprising a matrix of
PT oligonucleotides for hybridisation for identifying nucleic acid sequences
PT in samples.
XX
XX Claim 22; Fig 9; 76pp; English.
XX
XX AAV45901-V45974 are oligonucleotides which are used as biosensors in a
CC method for using a microchip to identify a sequence in a nucleic acid of
CC a sample. Such chips can be used to identifying nucleotide sequences in
CC samples such as air, water, soil, blood, cells, tissue, tissue culture
CC and food. The same microchip can be used for hybridisation for more than
CC 20-30 times, without any noticeable deterioration of the hybridisation
CC signal. Also, parallel hybridisation of nucleic acids in a sample to many
CC oligonucleotides on a microchip is possible, allowing replication and
CC standardisation
XX
SQ Sequence 10 BP; 2 A; 2 C; 4 G; 2 T; 0 U; 0 Other;
XX
Query Match 62.0%; Score 6.2; DB 2; Length 10;
Best Local Similarity 40.0%; Pred. No. 6.7e+05;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
QY 1 UGCDGCHMD 10
Db 1 TCGCGGTCAA 10
XX
RESULT 5
AA32311
ID AAX32311 standard; RNA; 10 BP.
XX
XX AAX32311;
AC
XX
XX 16-JUN-1999 (first entry)
DT
XX
XX Radioactivator isolate no: 7.
DE
XX
XX Transcriptional regulator; DNA-binding moiety; promoter; transcription;
KW initiation; elongation; gene expression; cancer; anemia; erythrocyte;
KW riboactivator; R10 library; ss.
XX
XX Synthetic.
OS
XX
XX WO9910487-A2.
PN
XX
XX 04-MAR-1999.
PD
XX
XX 26-AUG-1998; 98WO-US017691.
PF
XX
XX 27-AUG-1997; 97US-0056857P.
PR
XX
XX (HARD ) HARVARD COLLEGE.
PA (UYBO-) UNIV BOSTON.
XX
XX Jarrell KA, Saha S, Ptashne M;
PI
XX WPI; 1999-204663/17.
DR
XX
XX New RNA transcriptional regulators - used for modulating gene expression
PT in vitro or in vivo, e.g. as therapeutic agents for treating e.g. cancer

```

```

PT or anaemia.
XX
XX Example 2; Page 23; 57pp; English.
XX
XX The invention provides novel transcriptional regulators (TR) that are
CC comprised of RNA molecules. The TR comprises a DNA-binding moiety and an
CC RNA linked to the DNA binding moiety, where the RNA has TR activity.
CC Methods of identifying such RNA TRs are also provided. The RNA TRs alter
CC the rate and/or the extent of transcription from a promoter when they are
CC delivered to a site that is operationally linked to that promoter. The
CC TRs can affect transcription initiation, elongation, reinitiation,
CC termination and pausing. They can be used in e.g. bacterial cells, yeast
CC cells, mammalian cells, insect cells, plant cells, reptile cells,
CC celanorate cells, and protozoan cells. They can be used as agents for
CC controlling gene expression, e.g. to modulate gene expression in vivo in
CC order to alleviate or correct a disease state, e.g. cancer, anemia and
CC other disorders related to erythrocyte production. Sequences AAX32305-312
CC represent radioactivator isolate sequences obtained from a R10 library
XX
SQ Sequence 10 BP; 2 A; 2 C; 3 G; 0 T; 3 U; 0 Other;
XX
Query Match 62.0%; Score 6.2; DB 2; Length 10;
Best Local Similarity 50.0%; Pred. No. 6.7e+05;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
QY 1 UGCDGCHMD 10
Db 1 UGCDGCAUCA 10
XX
RESULT 6
AAX32304
ID AAX32304 standard; RNA; 10 BP.
XX
XX AAX32304;
AC
XX
XX 16-JUN-1999 (first entry)
DT
XX
XX Transcriptional regulatory RNA sequence.
DE
XX
XX Transcriptional regulator; DNA-binding moiety; promoter; transcription;
KW initiation; elongation; gene expression; cancer; anemia; erythrocyte; ss.
XX
XX Synthetic.
OS
XX
XX WO9910487-A2.
PN
XX
XX 04-MAR-1999.
PD
XX
XX 26-AUG-1998; 98WO-US017691.
PF
XX
XX 27-AUG-1997; 97US-0056857P.
PR
XX
XX (HARD ) HARVARD COLLEGE.
PA (UYBO-) UNIV BOSTON.
XX
XX Jarrell KA, Saha S, Ptashne M;
PI
XX WPI; 1999-204663/17.
DR
XX
XX New RNA transcriptional regulators - used for modulating gene expression
PT in vitro or in vivo, e.g. as therapeutic agents for treating e.g. cancer
PT or anaemia.
XX
XX Claim 9; Page 45; 57pp; English.
XX
XX The invention provides novel transcriptional regulators (TR) that are
CC comprised of RNA molecules. The TR comprises a DNA-binding moiety and an
CC RNA linked to the DNA binding moiety, where the RNA has TR activity.
CC Methods of identifying such RNA TRs are also provided. The RNA TRs alter
CC the rate and/or the extent of transcription from a promoter when they are
CC delivered to a site that is operationally linked to that promoter. The
CC TRs can affect transcription initiation, elongation, reinitiation,

```

CC termination and pausing. They can be used in e.g. bacterial cells, yeast  
 CC cells, mammalian cells, insect cells, plant cells, reptile cells,  
 CC ctenophore cells, and protozoan cells. They can be used as agents for  
 CC controlling gene expression, e.g. to modulate gene expression in vivo in  
 CC order to alleviate or correct a disease state, e.g. cancer, anemia and  
 CC other disorders related to erythrocyte production. The present sequence  
 CC represents a transcriptional regulatory RNA sequence

Sequence 10 BP; 0 A; 1 C; 3 G; 0 T; 1 U; 5 Other;

Query Match Best Local Similarity 62.0%; Score 6.2; DB 2; Length 10;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 UGCDGGHND 10  
 DB 1 UGCDGGHND 10

RESULT 7  
 AAX32307

ID AAX32307 standard; RNA; 10 BP.

AC AAX32307;

DT 16-JUN-1999 (first entry)

DE Radioactivator isolate no: 3.

KW Transcriptional regulator; DNA-binding moiety; promoter; transcription;  
 KW initiation; elongation; gene expression; cancer; anemia; erythrocyte;  
 KW riboactivator; R10 library; ss.

OS Synthetic.

PN WO9910487-A2.

PD 04-MAR-1999.

PF 26-AUG-1998; 98WO-US017691.

PR 27-AUG-1997; 97US-0056857P.

PA (HARD) HARVARD COLLEGE.

PA (UYBO-) UNIV BOSTON.

PI Jarrell KA, Saha S, Ptashne M;

DR WPI; 1999-204663/17.

PT New RNA transcriptional regulators - used for modulating gene expression  
 PT in vitro or in vivo, e.g. as therapeutic agents for treating e.g. cancer  
 PT or anaemia.

PS Example 2; Page 23; 57pp; English.

CC The invention provides novel transcriptional regulators (TR) that are  
 CC comprised of RNA molecules. The TR comprises a DNA-binding moiety and an  
 CC RNA linked to the DNA binding moiety, where the RNA has TR activity.  
 CC Methods of identifying such RNA TRs are also provided. The RNA TRs alter  
 CC the rate and/or the extent of transcription from a promoter when they are  
 CC delivered to a site that is operationally linked to that promoter. The  
 CC TRs can affect transcription initiation, elongation, reinitiation. The  
 CC termination and pausing. They can be used in e.g. bacterial cells, yeast  
 CC cells, mammalian cells, insect cells, plant cells, reptile cells,  
 CC ctenophore cells, and protozoan cells. They can be used as agents for  
 CC controlling gene expression, e.g. to modulate gene expression in vivo in  
 CC order to alleviate or correct a disease state, e.g. cancer, anemia and  
 CC other disorders related to erythrocyte production. Sequences AAX32305-312  
 CC represent radioactivator isolate sequences obtained from a R10 library

Sequence 10 BP; 1 A; 2 C; 4 G; 0 T; 3 U; 0 Other;

Query Match Best Local Similarity 50.0%; Score 6.2; DB 2; Length 10;  
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 UGCDGGHND 10  
 DB 1 UGCGGUCAU 10

RESULT 8  
 AAX32305

ID AAX32305 standard; RNA; 10 BP.

AC AAX32305;

DT 16-JUN-1999 (first entry)

DE Radioactivator isolate no: 1.

KW Transcriptional regulator; DNA-binding moiety; promoter; transcription;  
 KW initiation; elongation; gene expression; cancer; anemia; erythrocyte;  
 KW riboactivator; R10 library; ss.

OS Synthetic.

PN WO9910487-A2.

PD 04-MAR-1999.

PF 26-AUG-1998; 98WO-US017691.

PR 27-AUG-1997; 97US-0056857P.

PA (HARD) HARVARD COLLEGE.

PA (UYBO-) UNIV BOSTON.

PI Jarrell KA, Saha S, Ptashne M;

DR WPI; 1999-204663/17.

PT New RNA transcriptional regulators - used for modulating gene expression  
 PT in vitro or in vivo, e.g. as therapeutic agents for treating e.g. cancer  
 PT or anaemia.

PS Example 2; Page 23; 57pp; English.

CC The invention provides novel transcriptional regulators (TR) that are  
 CC comprised of RNA molecules. The TR comprises a DNA-binding moiety and an  
 CC RNA linked to the DNA binding moiety, where the RNA has TR activity.  
 CC Methods of identifying such RNA TRs are also provided. The RNA TRs alter  
 CC the rate and/or the extent of transcription from a promoter when they are  
 CC delivered to a site that is operationally linked to that promoter. The  
 CC TRs can affect transcription initiation, elongation, reinitiation. The  
 CC termination and pausing. They can be used in e.g. bacterial cells, yeast  
 CC cells, mammalian cells, insect cells, plant cells, reptile cells,  
 CC ctenophore cells, and protozoan cells. They can be used as agents for  
 CC controlling gene expression, e.g. to modulate gene expression in vivo in  
 CC order to alleviate or correct a disease state, e.g. cancer, anemia and  
 CC other disorders related to erythrocyte production. Sequences AAX32305-312  
 CC represent radioactivator isolate sequences obtained from a R10 library

Sequence 10 BP; 1 A; 2 C; 5 G; 0 T; 2 U; 0 Other;

Query Match Best Local Similarity 62.0%; Score 6.2; DB 2; Length 10;  
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 UGCDGGHND 10  
 DB 1 UGCGGUCAG 10

RESULT 9

AA32309  
ID AAX32309 standard; RNA; 10 BP.  
XX  
AC AAX32309;  
XX  
DT 16-JUN-1999 (first entry)  
XX  
DE Radioactivator isolate no: 5.  
XX  
KW Transcriptional regulator; DNA-binding moiety; promoter; transcription;  
KW initiation; elongation; gene expression; cancer; anemia; erythrocyte;  
KW riboactivator; R10 library; ss.  
XX  
OS Synthetic.  
XX  
PN WO9910487-A2.  
XX  
PD 04-MAR-1999.  
XX  
PF 26-AUG-1998; 98WO-US017691.  
XX  
PR 27-AUG-1997; 97US-0056857P.  
XX  
PA (HARD ) HARVARD COLLEGE.  
PA (UYBO-) UNIV BOSTON.  
XX  
PI Jarrell KA, Saha S, Ptashne M;  
XX  
DR WPI; 1999-204663/17.  
XX  
PT New RNA transcriptional regulators - used for modulating gene expression  
PT in vitro or in vivo, e.g. as therapeutic agents for treating e.g. cancer  
PT or anaemia.  
XX  
PS Example 2; Page 23; 57pp; English.  
XX  
CC The invention provides novel transcriptional regulators (TR) that are  
CC comprised of RNA molecules. The TR comprises a DNA-binding moiety and an  
CC RNA linked to the DNA binding moiety, where the RNA has TR activity.  
CC Methods of identifying such RNA TRs are also provided. The RNA TRs alter  
CC the rate and/or the extent of transcription from a promoter when they are  
CC delivered to a site that is operationally linked to that promoter. The  
CC TRs can affect transcription initiation, elongation, reinitiation,  
CC termination and pausing. They can be used in e.g. bacterial cells, yeast  
CC cells, mammalian cells, insect cells, plant cells, reptile cells,  
CC celanorate cells, and protozoan cells. They can be used as agents for  
CC controlling gene expression, e.g. to modulate gene expression in vivo in  
CC order to alleviate or correct a disease state, e.g. cancer, anemia and  
CC other disorders related to erythrocyte production. Sequences AAX32305-312  
CC represent radioactivator isolate sequences obtained from a R10 library  
XX  
SQ Sequence 10 BP; 2 A; 2 C; 4 G; 0 T; 2 U; 0 Other;  
XX  
Query Match 62.0%; Score 6.2; DB 2; Length 10;  
Best Local Similarity 50.0%; Pred. No. 6.7e+05;  
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
XX  
OY 1 UGCGGHNMD 10  
DB 1 UGCGGGAUCA 10  
XX  
RESULT 10  
AAX32310  
ID AAX32310 standard; RNA; 10 BP.  
XX  
AC AAX32310;  
XX  
DT 16-JUN-1999 (first entry)  
XX  
DE Radioactivator isolate no: 6.  
XX  
KW Transcriptional regulator; DNA-binding moiety; promoter; transcription;

KW initiation; elongation; gene expression; cancer; anemia; erythrocyte;  
KW riboactivator; R10 library; ss.  
XX  
OS Synthetic.  
XX  
PN WO9910487-A2.  
XX  
PD 04-MAR-1999.  
XX  
PF 26-AUG-1998; 98WO-US017691.  
XX  
PR 27-AUG-1997; 97US-0056857P.  
XX  
PA (HARD ) HARVARD COLLEGE.  
PA (UYBO-) UNIV BOSTON.  
XX  
PI Jarrell KA, Saha S, Ptashne M;  
XX  
DR WPI; 1999-204663/17.  
XX  
PT New RNA transcriptional regulators - used for modulating gene expression  
PT in vitro or in vivo, e.g. as therapeutic agents for treating e.g. cancer  
PT or anaemia.  
XX  
PS Example 2; Page 23; 57pp; English.  
XX  
CC The invention provides novel transcriptional regulators (TR) that are  
CC comprised of RNA molecules. The TR comprises a DNA-binding moiety and an  
CC RNA linked to the DNA binding moiety, where the RNA has TR activity.  
CC Methods of identifying such RNA TRs are also provided. The RNA TRs alter  
CC the rate and/or the extent of transcription from a promoter when they are  
CC delivered to a site that is operationally linked to that promoter. The  
CC TRs can affect transcription initiation, elongation, reinitiation,  
CC termination and pausing. They can be used in e.g. bacterial cells, yeast  
CC cells, mammalian cells, insect cells, plant cells, reptile cells,  
CC celanorate cells, and protozoan cells. They can be used as agents for  
CC controlling gene expression, e.g. to modulate gene expression in vivo in  
CC order to alleviate or correct a disease state, e.g. cancer, anemia and  
CC other disorders related to erythrocyte production. Sequences AAX32305-312  
CC represent radioactivator isolate sequences obtained from a R10 library  
XX  
SQ Sequence 10 BP; 1 A; 2 C; 4 G; 0 T; 3 U; 0 Other;  
XX  
Query Match 62.0%; Score 6.2; DB 2; Length 10;  
Best Local Similarity 50.0%; Pred. No. 6.7e+05;  
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
XX  
OY 1 UGCGGHNMD 10  
DB 1 UGCGGGAUCA 10  
XX  
RESULT 11  
AAX32308  
ID AAX32308 standard; RNA; 10 BP.  
XX  
AC AAX32308;  
XX  
DT 16-JUN-1999 (first entry)  
XX  
DE Radioactivator isolate no: 4.  
XX  
KW Transcriptional regulator; DNA-binding moiety; promoter; transcription;  
KW initiation; elongation; gene expression; cancer; anemia; erythrocyte;  
KW riboactivator; R10 library; ss.  
XX  
OS Synthetic.  
XX  
PN WO9910487-A2.  
XX  
PD 04-MAR-1999.  
XX  
PF 26-AUG-1998; 98WO-US017691.

```

XX 27-AUG-1997; 97US-0056857P.
XX
XX (HARD ) HARVARD COLLEGE.
XX (OTBO-) UNIV BOSTON.
XX
XX Jarrell KA, Saha S, Ptashne M;
XX WPI; 1999-204663/17.
XX
XX New RNA transcriptional regulators - used for modulating gene expression
XX in vitro or in vivo, e.g. as therapeutic agents for treating e.g. cancer
XX or anaemia.
XX
XX Example 2; Page 23; 57pp; English.
XX
XX The invention provides novel transcriptional regulators (TR) that are
XX comprised of RNA molecules. The TR comprises a DNA-binding moiety and an
XX RNA linked to the DNA binding moiety, where the RNA has TR activity.
XX Methods of identifying such RNA TRs are also provided. The RNA TRs alter
XX the rate and/or the extent of transcription from a promoter when they are
XX delivered to a site that is operationally linked to that promoter. The
XX TRs can affect transcription initiation, elongation, reinitiation,
XX termination and pausing. They can be used in e.g. bacterial cells, yeast
XX cells, mammalian cells, insect cells, plant cells, reptile cells,
XX celomorate cells, and protozoan cells. They can be used as agents for
XX controlling gene expression, e.g. to modulate gene expression in vivo in
XX order to alleviate or correct a disease state, e.g. cancer, anaemia and
XX other disorders related to erythrocyte production. Sequences AA33205-312
XX represent radioactivator isolate sequences obtained from a R10 library
XX
XX Sequence 10 BP; 0 A; 2 C; 5 G; 0 T; 3 U; 0 Other;
XX
XX Query Match 62.0%; Score 6.2; DB 2; Length 10;
XX Best Local Similarity 50.0%; Pred. No. 6.7e+05;
XX Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 UCCDGGHMD 10
XX |||:::
XX 1 UCCGGGUNC 10
XX
XX RESULT 12
XX AA277884/C
XX ID AA277884 standard; DNA; 10 BP.
XX
XX AA277884;
XX
XX 10-APR-2000 (first entry)
XX
XX Human dendritic cell SAGE tag, SEQ ID NO:312.
XX
XX SAGE tag; serial analysis of gene expression; antigen-presenting cell;
XX APC; monocyte-derived dendritic cell; differential gene expression;
XX immunostimulatory cofactor; costimulatory factor; CTU;
XX cytotoxic T-lymphocyte; tumour antigen; immunotherapy; anticancer; ss.
XX
XX Homo sapiens.
XX
XX MO9965924-A2.
XX
XX 23-DEC-1999.
XX
XX 18-JUN-1999; 99MO-US013800.
XX
XX 19-JUN-1998; 98US-0089833P.
XX 19-JUN-1998; 98US-0089844P.
XX 19-JUN-1998; 98US-0089853P.
XX 19-JUN-1998; 98US-0089878P.
XX 19-JUN-1998; 98US-0089991P.
XX 19-JUN-1998; 98US-0089992P.
XX 19-JUN-1998; 98US-0089993P.
XX 19-JUN-1998; 98US-0089994P.

```

```

PR 19-JUN-1998; 98US-0089997P.
PR 19-JUN-1998; 98US-0089999P.
PR 19-JUN-1998; 98US-0090000P.
PR 19-JUN-1998; 98US-0090003P.
PR 19-JUN-1998; 98US-0090035P.
PR 19-JUN-1998; 98US-0090036P.
PR 19-JUN-1998; 98US-0090039P.
PR 19-JUN-1998; 98US-0090041P.
PR 19-JUN-1998; 98US-0090042P.
PR 19-JUN-1998; 98US-0090043P.
PR 19-JUN-1998; 98US-0090044P.
PR 19-JUN-1998; 98US-0090045P.
PR 19-JUN-1998; 98US-0090047P.
PR 19-JUN-1998; 98US-0090048P.
PR 19-JUN-1998; 98US-0090072P.
PR 19-JUN-1998; 98US-0090076P.
PR 19-JUN-1998; 98US-0090077P.
PR 19-JUN-1998; 98US-0090078P.
PR 19-JUN-1998; 98US-0090079P.
PR 19-JUN-1998; 98US-0090080P.
PR 08-DEC-1998; 98US-0111715P.
XX
XX (GENZ ) GENZYME CORP.
XX (ROBE/) ROBERTS B. L.
XX (SHAN/) SHANKARA S.
XX
XX Roberts BL, Shankara S;
XX WPI; 2000-106077/09.
XX
XX Isolated polynucleotides differentially expressed in antigen-presenting
XX cells, useful in gene vaccines against cancer.
XX
XX Claim 1; Page 73; 130pp; English.
XX
XX Sequences AA27573-279709 represent SAGE (serial analysis of gene
XX expression) tags used to identify mRNA transcripts encoding
XX immunostimulatory cofactor proteins which are preferentially or
XX differentially expressed in monocyte-derived dendritic cells compared
XX with monocytes. Some of the transcripts correspond to known genes or ESTs
XX (expressed sequence tags) which were previously unknown to be
XX preferentially or differentially expressed in dendritic cells, while
XX other transcripts correspond to novel genes. Antigen-presenting cell
XX (APC)-associated costimulatory factors play an important role in the
XX activation of the cytotoxic immune response, particularly against tumour
XX cells. Tumour antigen presentation via the MHC (major histocompatibility
XX complex) and subsequent recognition by T-cell receptors is alone
XX insufficient to activate a robust cytotoxic immune response that can lyse
XX the tumour cells, immunostimulatory cofactors also being required for
XX efficient activation of cytotoxic T-lymphocytes (CTLs). Nucleic acid
XX sequences identified using the SAGE tags have several potential uses.
XX They may be used in vaccines to induce an immune response, particularly
XX against a tumour antigen; to modulate the genotype of an APC; to screen
XX for agents that modulate expression of differentially expressed genes in
XX an APC; and as hybridisation probes/amplification primers for the
XX diagnosis, prognosis and monitoring of diseases related to abnormal
XX expression of these genes. Detection of the dendritic cell differentially
XX expressed genes, or of their encoded proteins, can be used to identify
XX cells as belonging to the monocyte lineage. Cells containing these genes
XX can be used in active immunotherapy (or to stimulate production of a
XX population of antigen-specific effector cells) and vectors containing
XX them are used in gene therapy. Co-administration of tumour antigens and
XX APC-associated costimulatory factors ensures adequate antigen
XX presentation to endogenous APCs and upregulates the APCs for the
XX secretion of co-stimulatory signals, migration to T cell-rich sites,
XX recruitment of T cell growth factors and secretion of chemokines for
XX recruitment of immune effector cells
XX
XX Sequence 10 BP; 1 A; 3 C; 1 G; 5 T; 0 U; 0 Other;
XX
XX Query Match 62.0%; Score 6.2; DB 3; Length 10;
XX Best Local Similarity 40.0%; Pred. No. 6.7e+05;
XX Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

```



PA (SHAN/) SHANKARA S.  
 XX  
 PI Roberts BL, Shankara S;  
 XX  
 DR WPI: 2000-106079/09.  
 XX  
 PT Isolated polynucleotides differentially expressed between metastatic and  
 PT non-metastatic breast cancer cells, useful for diagnosis, prevention and  
 PT treatment of cancer.  
 XX  
 PS Claim 1, Page 65, 219pp; English.

CC AA280767 to AA283941 represent tags corresponding to distinct transcripts  
 CC that are preferentially transcribed in the metastatic breast tumour  
 CC tissue (i.e. are upregulated in metastatic breast tumour cells). AA283942  
 CC to AA286677 represent tags corresponding to distinct transcripts that are  
 CC preferentially transcribed in the primary or non-metastatic breast tumour  
 CC tissue (i.e. are downregulated in metastatic breast tumour cells). These  
 CC transcripts can be used for diagnosis, prognosis, monitoring and  
 CC treatment of breast cancer, particularly where metastatic. Diagnosis is  
 CC by standard immunoassays or hybridisation/amplification reactions.  
 CC Compounds that modulate expression of the transcripts are potentially  
 CC useful for treatment of (metastatic) breast cancer, while promoters from  
 CC the transcripts are used to direct expression, in selected cell types, of  
 CC e.g. therapeutic genes (also ribozymes or antisense sequences),  
 CC particularly an antigen-encoding sequence for use in gene or cell-based  
 CC vaccines. Polypeptides encoded by the transcripts are also useful in  
 CC vaccines; for diagnosing breast cancer and for raising specific  
 CC antibodies (Ab). Ab are used to detect the polypeptides or as therapeutic  
 CC agents. Host cells that produce the polypeptides can be used to expand  
 CC and isolate populations of educated, antigen-specific immune effector  
 CC cells, e.g. cytotoxic T lymphocytes, and these used for adoptive  
 CC immunotherapy  
 XX  
 SQ Sequence 10 BP; 2 A; 2 C; 3 G; 3 T; 0 U; 0 Other;

Query Match 62.0%; Score 6.2; DB 3; Length 10;  
 Best Local Similarity 40.0%; Pred. No. 6.7e+05;  
 Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 UGCDGGMND 10  
 Db 1 TGCAAGTACT 10

## RESULT 15

AA285761  
 ID AA285761 standard; DNA; 10 BP.

XX  
 AC AA285761;

DT 07-APR-2000 (first entry)

DE Metastatic breast tumour cell downregulated transcript tag #4995.

XX  
 KW Human; metastatic breast tumour tissue; breast cancer; tag; primer;  
 KW non-metastatic breast tumour tissue; gene therapy; anticancer;  
 XX antimetastatic; vaccine; diagnosis; ss.

OS Homo sapiens.

XX  
 PN MO9965928-A2.

XX  
 PD 23-DEC-1999.

XX  
 PF 18-JUN-1999; 99WO-US013647.

XX  
 PR 19-JUN-1998; 98US-0089853P.

XX  
 PR 19-JUN-1998; 98US-008997P.

XX  
 PR 19-JUN-1998; 98US-0090039P.

XX  
 PR 19-JUN-1998; 98US-0090040P.

XX  
 PR 19-JUN-1998; 98US-0090041P.

PA (GENZ ) GENZYME CORP.  
 PA (ROBE/) ROBERTS B L.  
 PA (SHAN/) SHANKARA S.  
 XX  
 PI Roberts BL, Shankara S;  
 XX  
 DR WPI: 2000-106079/09.  
 XX  
 PT Isolated polynucleotides differentially expressed between metastatic and  
 PT non-metastatic breast cancer cells, useful for diagnosis, prevention and  
 PT treatment of cancer.  
 XX  
 PS Claim 1, Page 192, 219pp; English.

CC AA280767 to AA283941 represent tags corresponding to distinct transcripts  
 CC that are preferentially transcribed in the metastatic breast tumour  
 CC tissue (i.e. are upregulated in metastatic breast tumour cells). AA283942  
 CC to AA286677 represent tags corresponding to distinct transcripts that are  
 CC preferentially transcribed in the primary or non-metastatic breast tumour  
 CC tissue (i.e. are downregulated in metastatic breast tumour cells). These  
 CC transcripts can be used for diagnosis, prognosis, monitoring and  
 CC treatment of breast cancer, particularly where metastatic. Diagnosis is  
 CC by standard immunoassays or hybridisation/amplification reactions.  
 CC Compounds that modulate expression of the transcripts are potentially  
 CC useful for treatment of (metastatic) breast cancer, while promoters from  
 CC the transcripts are used to direct expression, in selected cell types, of  
 CC e.g. therapeutic genes (also ribozymes or antisense sequences),  
 CC particularly an antigen-encoding sequence for use in gene or cell-based  
 CC vaccines. Polypeptides encoded by the transcripts are also useful in  
 CC vaccines; for diagnosing breast cancer and for raising specific  
 CC antibodies (Ab). Ab are used to detect the polypeptides or as therapeutic  
 CC agents. Host cells that produce the polypeptides can be used to expand  
 CC and isolate populations of educated, antigen-specific immune effector  
 CC cells, e.g. cytotoxic T lymphocytes, and these used for adoptive  
 CC immunotherapy  
 XX  
 SQ Sequence 10 BP; 3 A; 1 C; 4 G; 2 T; 0 U; 0 Other;

Query Match 62.0%; Score 6.2; DB 3; Length 10;  
 Best Local Similarity 40.0%; Pred. No. 6.7e+05;  
 Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 UGCDGGMND 10  
 Db 1 TGCTGGAGAA 10

Search completed: October 19, 2004, 15:04:54  
 Job time : 293 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 19, 2004, 14:46:35 ; Search time 57 Seconds  
(without alignments)  
124.700 Million cell updates/sec

Title: US-09-762-648-4  
Perfect score: 10  
Sequence: 1 UGCDGHNMD 10

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	6.2	62.0	10	US-08-388-353-70	Sequence 70, Appl
C 2	6.2	62.0	10	US-08-388-353-504	Sequence 504, App
C 3	6.2	62.0	10	US-08-488-551B-70	Sequence 70, Appl
C 4	6.2	62.0	10	US-08-488-551B-504	Sequence 504, App
C 5	6.2	62.0	10	US-08-488-551B-822	Sequence 822, App
C 6	6.2	62.0	10	US-08-522-384-102	Sequence 102, App
C 7	6.2	62.0	10	US-09-261-115-71	Sequence 71, Appl
C 8	6.2	62.0	10	US-09-508-753B-22	Sequence 22, Appl
C 9	6.2	62.0	11	US-08-481-658B-75	Sequence 75, Appl
C 10	6.2	62.0	11	US-08-477-504A-75	Sequence 75, Appl
C 11	6.2	62.0	11	US-08-486-756A-75	Sequence 75, Appl
C 12	6.2	62.0	11	US-08-485-862B-75	Sequence 75, Appl
C 13	6.2	62.0	11	US-08-787-739-75	Sequence 99, Appl
C 14	6.2	62.0	11	US-08-487-077A-75	Sequence 75, Appl
C 15	6.2	62.0	11	US-08-485-863A-75	Sequence 75, Appl
C 16	6.2	62.0	11	US-08-485-049D-75	Sequence 75, Appl
C 17	6.2	62.0	11	US-08-482-918-99	Sequence 99, Appl
C 18	6.2	62.0	11	US-09-224-681-99	Sequence 99, Appl
C 19	6.2	62.0	11	US-08-336-728A-99	Sequence 99, Appl
C 20	6.2	62.0	11	US-09-178-115-75	Sequence 75, Appl
C 21	6.2	62.0	11	US-09-177-776-75	Sequence 75, Appl
C 22	6.2	62.0	11	US-09-249-155A-303	Sequence 303, App
C 23	6.2	62.0	11	US-09-635-251-99	Sequence 99, Appl
C 24	6.2	62.0	11	US-09-772-719B-75	Sequence 75, Appl
C 25	6.2	62.0	12	US-08-035-928-18	Sequence 18, Appl
C 26	6.2	62.0	12	US-08-441-887A-177	Sequence 177, App
C 27	6.2	62.0	12	US-08-993-118-6	Sequence 6, Appl

C 28	6.2	62.0	12	US-08-845-528C-6	Sequence 6, Appl
C 29	6.2	62.0	12	US-09-620-926-1	Sequence 1, Appl
C 30	6.2	62.0	12	US-09-528-404-7	Sequence 7, Appl
C 31	6.2	62.0	12	US-09-066-281B-6	Sequence 6, Appl
C 32	6.2	62.0	12	US-09-574-117A-9	Sequence 9, Appl
C 33	6.2	62.0	12	US-09-468-433C-6	Sequence 6, Appl
C 34	6.2	62.0	13	5196516-5	Patent No. 5196516
C 35	6.2	62.0	14	US-08-332-747-9	Sequence 9, Appl
C 36	6.2	62.0	14	US-08-332-747-17	Sequence 17, Appl
C 37	6.2	62.0	14	US-08-998-099-339	Sequence 339, App
C 38	6.2	62.0	14	US-08-998-099-368	Sequence 368, App
C 39	6.2	62.0	14	US-09-475-947A-278	Sequence 278, App
C 40	6.2	62.0	15	US-08-319-492B-413	Sequence 413, App
C 41	6.2	62.0	15	US-08-363-240A-36	Sequence 36, Appl
C 42	6.2	62.0	15	US-08-363-240A-69	Sequence 69, Appl
C 43	6.2	62.0	15	US-08-363-240A-543	Sequence 543, App
C 44	6.2	62.0	15	US-08-363-240A-592	Sequence 592, App
C 45	6.2	62.0	15	US-08-363-240A-751	Sequence 751, App

## ALIGNMENTS

RESULT 1  
US-08-388-353-70/c  
Sequence 70, Application US/08388353  
Patent No. 6010895  
GENERAL INFORMATION:  
APPLICANT: Deacon, Nicholas J.  
APPLICANT: Learmont, Jennifer C.  
APPLICANT: McPhee, Dale A.  
APPLICANT: Crowe, Suzanne  
APPLICANT: Cooper, David  
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1  
NUMBER OF SEQUENCES: 800  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Scully, Scott, Murphy & Presser  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: New York  
COUNTRY: United States  
ZIP: 11530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/388,353  
FILING DATE: 14-FEB-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Digiglio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 9606  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
TELEX: 230 901 SAMS UR  
INFORMATION FOR SEQ ID NO: 70:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-388-353-70  
Query Match 62.0%; Score 6.2; DB 3; Length 10;  
Best Local Similarity 40.0%; Pred. No. 8.9e+04;  
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 UGCDGHNMD 10

Db 10 TGCTGCTCA 1

```

RESULT 2
US-08-388-353-504/c
; Sequence 504, Application US/08388353
; Patent No. 6010895
; GENERAL INFORMATION:
; APPLICANT: Deacon, Nicholas J.
; APPLICANT: Learmont, Jennifer C.
; APPLICANT: McPhee, Dale A.
; APPLICANT: Crowe, Suzanne
; APPLICANT: Cooper, David
; TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
; NUMBER OF SEQUENCES: 800
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/388,353
; FILING DATE: 14-FEB-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9606
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELE: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 504:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-388-353-504

```

Query Match 62.0%; Score 6.2; DB 3; Length 10;  
 Best Local Similarity 40.0%; Pred. No. 8.9e+04;  
 Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

```

Oy 1 UGCDGHNMD 10
Db 10 TGCTGCTCA 1

RESULT 3
US-08-488-551B-70/c
; Sequence 70, Application US/08488551B
; Patent No. 6015661
; GENERAL INFORMATION:
; APPLICANT: Nicholas J. Deacon
; APPLICANT: Dale A. McPhee
; APPLICANT: David Cooper
; TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
; NUMBER OF SEQUENCES: 841
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 GARDEN CITY PLAZA
; CITY: GARDEN CITY
; STATE: NEW YORK

```

```

; COUNTRY: U.S.A.
; ZIP: 11530-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,551B
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PM3864 (AU)
; FILING DATE: 14-FEB-1994
; APPLICATION NUMBER: PM4002 (AU)
; FILING DATE: 21-FEB-1994
; APPLICATION NUMBER: PM0284 (AU)
; FILING DATE: 23-DEC-1994
; APPLICATION NUMBER: US 08/388,353
; FILING DATE: 14-FEB-1995
; APPLICATION NUMBER: PM3021/95
; FILING DATE: 17-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: FRANK S. DIGIGLIO
; REFERENCE/DOCKET NUMBER: 96062
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; INFORMATION FOR SEQ ID NO: 70:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-488-551B-70

```

Query Match 62.0%; Score 6.2; DB 3; Length 10;  
 Best Local Similarity 40.0%; Pred. No. 8.9e+04;  
 Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

```

Oy 1 UGCDGHNMD 10
Db 10 TGCTGCTCA 1

RESULT 4
US-08-488-551B-504/c
; Sequence 504, Application US/08488551B
; Patent No. 6015661
; GENERAL INFORMATION:
; APPLICANT: Nicholas J. Deacon
; APPLICANT: Dale A. McPhee
; APPLICANT: David Cooper
; TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
; NUMBER OF SEQUENCES: 841
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 GARDEN CITY PLAZA
; CITY: GARDEN CITY
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 11530-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,551B
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PM3864 (AU)
; FILING DATE: 14-FEB-1994

```

```

APPLICATION NUMBER: PM4002 (AU)
FILING DATE: 21-FEB-1994
APPLICATION NUMBER: PN0284 (AU)
FILING DATE: 23-DEC-1994
APPLICATION NUMBER: US 08/388,353
FILING DATE: 14-FEB-1995
APPLICATION NUMBER: PN3021/95
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGIGLIO
REFERENCE/DOCKET NUMBER: 96062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
INFORMATION FOR SEQ ID NO: 504:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-488-551B-504

Query Match
Best Local Similarity 62.0%; Score 6.2; DB 3; Length 10;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 UGCDGSHNMD 10
Db 10 TGCAGCTCA 1

RESULT 5
US-08-488-551B-822/c
Sequence 822, Application US/08488551B
Patent No. 6015661
GENERAL INFORMATION:
APPLICANT: Nicholas J. Deacon
APPLICANT: Dale A. McPhee
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 841
CORRESPONDENCE ADDRESS:
ADDRESSER: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,551B
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PM3864 (AU)
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: PM4002 (AU)
FILING DATE: 21-FEB-1994
APPLICATION NUMBER: PN0284 (AU)
FILING DATE: 23-DEC-1994
APPLICATION NUMBER: US 08/388,353
FILING DATE: 14-FEB-1995
APPLICATION NUMBER: PN3021/95
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGIGLIO
REFERENCE/DOCKET NUMBER: 96062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343

```

```

TELEFAX: (516) 742-4366
INFORMATION FOR SEQ ID NO: 822:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-488-551B-822

Query Match
Best Local Similarity 62.0%; Score 6.2; DB 3; Length 10;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 UGCDGSHNMD 10
Db 10 TGCAGCTCA 1

RESULT 6
US-08-522-384-102
Sequence 102, Application US/08522384
Patent No. 6110667
GENERAL INFORMATION:
APPLICANT: LOPEZ-NIETO, CARLOS E
APPLICANT: NIGAM, SANJAY KUMAR
TITLE OF INVENTION: PROCESSES, APPARATUS AND COMPOSITIONS FOR
FILE REFERENCE: 2458-4029
CURRENT APPLICATION NUMBER: US/08/522,384
CURRENT FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 122
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 102
LENGTH: 10
TYPE: DNA
ORGANISM: Unknown Organism
FEATURE:
OTHER INFORMATION: Description of Unknown Organism: Primer
US-08-522-384-102

Query Match
Best Local Similarity 62.0%; Score 6.2; DB 3; Length 10;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 UGCDGSHNMD 10
Db 1 TGCCTGCTCAT 10

RESULT 7
US-09-261-115-71
Sequence 71, Application US/09261115
Patent No. 6458584
GENERAL INFORMATION:
APPLICANT: MIRZABEKOV, ANDREI
APPLICANT: GUSCHIN, DMITRY Y.
APPLICANT: SHIK, VALENTINE
APPLICANT: DROBYSHEV, ALEKSEI
APPLICANT: FOTIN, ALEXANDER
APPLICANT: YERSHOV, GENNADIY
APPLICANT: LYISOV, YU
TITLE OF INVENTION: CUSTOMIZED OLIGONUCLEOTIDE MICROCHIPS THAT CONVERT
TITLE OF INVENTION: MULTIPLE GENETIC INFORMATION TO SIMPLE PATTERNS, ARE
FILE REFERENCE: 21416/90184
CURRENT APPLICATION NUMBER: US/09/261,115
CURRENT FILING DATE: 1999-03-03
NUMBER OF SEQ ID NOS: 78
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 71
LENGTH: 10
TYPE: DNA

```

ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Customized  
OTHER INFORMATION: oligonucleotide  
US-09-261-115-71

Query Match 62.0%; Score 6.2; DB 4; Length 10;  
Best Local Similarity 40.0%; Pred. No. 8.9e+04;  
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 UGCDGGHND 10  
:||||: :  
Db 1 TGCGGTCAA 10

RESULT 8  
US-09-508-753B-22  
Sequence 22, Application US/09508753B  
Patent No. 6544736  
GENERAL INFORMATION:  
APPLICANT: Akira SHIMAMOTO  
APPLICANT: Yasuhito FURUICHI  
APPLICANT: Yuko SHIBATA  
APPLICANT: Hiroko FUNAKI  
APPLICANT: Eiji OHARA  
APPLICANT: Masamori WATAHAKI  
TITLE OF INVENTION: Method for synthesizing cDNA from mRNA sample  
FILE REFERENCE: 00162/HG  
CURRENT APPLICATION NUMBER: US/09/508,753B  
CURRENT FILING DATE: 2000-06-16  
PRIOR APPLICATION NUMBER: JP 9/270324  
PRIOR FILING DATE: 1997-09-18  
NUMBER OF SEQ ID NOS: 472  
SEQ ID NO 22  
LENGTH: 10  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Primer  
US-09-508-753B-22

Query Match 62.0%; Score 6.2; DB 4; Length 10;  
Best Local Similarity 40.0%; Pred. No. 8.9e+04;  
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 UGCDGGHND 10  
:||||: :  
Db 1 TGCTGGAAG 10

RESULT 9  
US-08-481-658B-75  
Sequence 75, Application US/08481658B  
Patent No. 5955075  
GENERAL INFORMATION:  
APPLICANT: Zavada, Jan  
APPLICANT: Pastorek, Jaromir  
APPLICANT: Pastorek, Silvia  
TITLE OF INVENTION: MN Gene and Protein  
NUMBER OF SEQUENCES: 86  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Leona L. Lauder  
STREET: 6 Mariposa Court  
CITY: Tiburon  
STATE: California  
COUNTRY: USA  
ZIP: 94920  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/481,658B  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/260,190  
FILING DATE: 15-JUN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Lauder, Leona L.  
REGISTRATION NUMBER: 30,863  
REFERENCE/DOCKET NUMBER: D-0021.3E  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-435-2034  
TELEFAX: 415-435-0727  
INFORMATION FOR SEQ ID NO: 75:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
DESCRIPTION: 5' donor consensus splice sequence  
US-08-481-658B-75

Query Match 62.0%; Score 6.2; DB 2; Length 11;  
Best Local Similarity 40.0%; Pred. No. 8.9e+04;  
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 UGCDGGHND 10  
:||||: :  
Db 1 TGCTGTGAG 10

RESULT 10  
US-08-477-504A-75  
Sequence 75, Application US/08477504A  
Patent No. 5972353  
GENERAL INFORMATION:  
APPLICANT: Zavada, Jan  
APPLICANT: Pastorek, Jaromir  
APPLICANT: Pastorek, Silvia  
TITLE OF INVENTION: MN Gene and Protein  
NUMBER OF SEQUENCES: 86  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Leona L. Lauder  
STREET: 6 Mariposa Court  
CITY: Tiburon  
STATE: California  
COUNTRY: USA  
ZIP: 94920  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,504A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/260,190  
FILING DATE: 15-JUN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Lauder, Leona L.  
REGISTRATION NUMBER: 30,863  
REFERENCE/DOCKET NUMBER: D-0021.3D  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-435-2034  
TELEFAX: 415-435-0727  
INFORMATION FOR SEQ ID NO: 75:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single

TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
DESCRIPTION: 5' donor consensus splice sequence  
US-08-477-504A-75

Query Match 62.0%; Score 6.2; DB 2; Length 11;  
Best Local Similarity 40.0%; Pred. No. 8.9e+04;  
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 UGCDGSHNMD 10  
Db 1 TGCTGTGTAG 10

RESULT 11  
US-08-486-756A-75  
Sequence 75, Application US/08486756A  
Patent No. 5981711  
GENERAL INFORMATION:  
APPLICANT: Zavada, Jan  
APPLICANT: Pastorekova, Silvia  
APPLICANT: Pastorek, Jaromir  
TITLE OF INVENTION: MN Gene and Protein  
NUMBER OF SEQUENCES: 86  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Leona L. Lauder  
STREET: 6 Mariposa Court  
CITY: Tiburon  
STATE: California  
COUNTRY: USA  
ZIP: 94920  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/486,756A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/260,190  
FILING DATE: 15-JUN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Lauder, Leona L.  
REGISTRATION NUMBER: 30,863  
REFERENCE/DOCKET NUMBER: D-0021.3C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-435-2034  
TELEFAX: 415-435-0727  
INFORMATION FOR SEQ ID NO: 75:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
DESCRIPTION: 5' donor consensus splice sequence  
US-08-486-756A-75

Query Match 62.0%; Score 6.2; DB 2; Length 11;  
Best Local Similarity 40.0%; Pred. No. 8.9e+04;  
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 UGCDGSHNMD 10  
Db 1 TGCTGTGTAG 10

RESULT 12  
US-08-485-862B-75  
Sequence 75, Application US/08485862B  
Patent No. 5989838

GENERAL INFORMATION:  
APPLICANT: Zavada, Jan  
APPLICANT: Pastorekova, Silvia  
APPLICANT: Pastorek, Jaromir  
TITLE OF INVENTION: MN Gene and Protein  
NUMBER OF SEQUENCES: 86  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Leona L. Lauder  
STREET: 6 Mariposa Court  
CITY: Tiburon  
STATE: California  
COUNTRY: USA  
ZIP: 94920  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/485,862B  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/477,504  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/260,190  
FILING DATE: 15-JUN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Lauder, Leona L.  
REGISTRATION NUMBER: 30,863  
REFERENCE/DOCKET NUMBER: D-0021.3D  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-435-2034  
TELEFAX: 415-435-0727  
INFORMATION FOR SEQ ID NO: 75:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
DESCRIPTION: 5' donor consensus splice sequence  
US-08-485-862B-75

Query Match 62.0%; Score 6.2; DB 2; Length 11;  
Best Local Similarity 40.0%; Pred. No. 8.9e+04;  
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 UGCDGSHNMD 10  
Db 1 TGCTGTGTAG 10

RESULT 13  
US-08-787-739-75  
Sequence 75, Application US/08787739  
Patent No. 6027887  
GENERAL INFORMATION:  
APPLICANT: Zavada, Jan  
APPLICANT: Pastorekova, Silvia  
APPLICANT: Pastorek, Jaromir  
TITLE OF INVENTION: MN Gene and Protein  
NUMBER OF SEQUENCES: 96  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Leona L. Lauder  
STREET: 369 Pine Street, Suite 610  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

```
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/787,739
FILING DATE: 24-JUN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,049
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/486,756
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/477,504
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/481,658
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,862
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,863
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/487,077
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.4
TELEPHONE: 415-981-2034
TELEFAX: 415-981-0332
INFORMATION FOR SEQ ID NO: 75:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
DESCRIPTION: 5' donor consensus splice sequence
US-08-787-739-75

Query Match      62.0%; Score 6.2; DB 3; Length 11;
Best Local Similarity 40.0%; Pred. No. 8.9e+04;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Cy      1 UGCDGHNMD 10
Db      1 TGCTGTGAG 10

RESULT 14
US-08-487-077A-75
Sequence 75, Application US/08487077A
Patent No. 6069242
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 6 Mariposa Court
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,077A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3H
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 75:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
DESCRIPTION: 5' donor consensus splice sequence
US-08-487-077A-75
```

```
Query Match      62.0%; Score 6.2; DB 3; Length 11;
Best Local Similarity 40.0%; Pred. No. 8.9e+04;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Cy      1 UGCDGHNMD 10
Db      1 TGCTGTGAG 10
```

```
RESULT 15
US-08-485-863A-75
Sequence 75, Application US/08485863A
Patent No. 6093548
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 6 Mariposa Court
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,863A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3G
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 75:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 base pairs
```

```

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; DESCRIPTION: 5' donor consensus splice sequence
US-08-485-863A-75

```

```

Query Match          62.0%; Score 6.2; DB 3; Length 11;
Best Local Similarity 40.0%; Pred. No. 8.9e+04;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      1 UGCDGGHMD 10
        :|||:|:|:|:
Db      1 TGCTGTGAG 10

```

Search completed: October 19, 2004, 16:08:36  
Job time : 58 secs

This Page Blank (uspto)

BEST AVAILABLE COPY



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 19, 2004, 15:00:01 : Search time 260 Seconds  
(without alignments)  
196.753 Million cell updates/sec

Title: US-09-762-648-4  
Perfect score: 10  
Sequence: 1 UGCDGSHMND 10

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 3403857 seqs, 255783690 residues

Total number of hits satisfying chosen parameters: 6807714

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 19: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 20: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	6.2	62.0	10 13 US-10-033-145-59	Sequence 59, App1
2	6.2	62.0	10 13 US-10-033-145-312	Sequence 312, App1
3	6.2	62.0	10 13 US-10-084-700-17	Sequence 17, App1
4	6.2	62.0	10 15 US-10-212-476-71	Sequence 71, App1
5	6.2	62.0	11 9 US-09-005-243-99	Sequence 99, App1
6	6.2	62.0	11 9 US-09-224-683-99	Sequence 99, App1
7	6.2	62.0	11 9 US-09-772-719-75	Sequence 75, App1
8	6.2	62.0	11 10 US-09-918-715-162	Sequence 162, App1
9	6.2	62.0	11 10 US-09-967-237-75	Sequence 75, App1
10	6.2	62.0	11 15 US-10-314-322-303	Sequence 303, App1
11	6.2	62.0	11 17 US-10-203-969A-5	Sequence 5, App1
12	6.2	62.0	11 17 US-10-450-797-357	Sequence 357, App1
13	6.2	62.0	11 17 US-10-450-797-421	Sequence 421, App1

C 14	6.2	62.0	11 17 US-10-450-797-663	Sequence 663, App1
C 15	6.2	62.0	11 17 US-10-175-608-99	Sequence 99, App1
C 16	6.2	62.0	12 13 US-10-085-108-6	Sequence 6, App1
C 17	6.2	62.0	12 15 US-10-076-047A-110	Sequence 110, App1
C 18	6.2	62.0	12 15 US-10-160-237-6	Sequence 6, App1
C 19	6.2	62.0	13 10 US-09-823-181-4	Sequence 4, App1
C 20	6.2	62.0	14 8 US-08-980-068B-12	Sequence 12, App1
C 21	6.2	62.0	14 8 US-08-980-068B-15	Sequence 15, App1
C 22	6.2	62.0	14 8 US-08-980-068B-17	Sequence 17, App1
C 23	6.2	62.0	14 10 US-09-880-313A-131	Sequence 131, App1
C 24	6.2	62.0	14 15 US-10-084-839-3065	Sequence 3065, App1
C 25	6.2	62.0	14 15 US-10-084-839-3069	Sequence 3069, App1
C 26	6.2	62.0	14 15 US-10-091-281-346	Sequence 346, App1
C 27	6.2	62.0	14 17 US-10-318-190-11	Sequence 11, App1
C 28	6.2	62.0	15 9 US-09-504-231A-1160	Sequence 1160, App1
C 29	6.2	62.0	15 9 US-09-504-231A-1161	Sequence 1161, App1
C 30	6.2	62.0	15 9 US-09-152-059-11	Sequence 11, App1
C 31	6.2	62.0	15 9 US-09-152-059-13	Sequence 13, App1
C 32	6.2	62.0	15 9 US-09-152-059-20	Sequence 20, App1
C 33	6.2	62.0	15 9 US-09-152-059-21	Sequence 21, App1
C 34	6.2	62.0	15 9 US-09-274-553D-1160	Sequence 1160, App1
C 35	6.2	62.0	15 9 US-09-274-553D-1161	Sequence 1161, App1
C 36	6.2	62.0	15 9 US-09-544-934B-106	Sequence 106, App1
C 37	6.2	62.0	15 9 US-09-263-959-59	Sequence 59, App1
C 38	6.2	62.0	15 9 US-09-263-959-69	Sequence 69, App1
C 39	6.2	62.0	15 9 US-09-263-959-83	Sequence 83, App1
C 40	6.2	62.0	15 10 US-09-771-933-175	Sequence 175, App1
C 41	6.2	62.0	15 10 US-09-880-313A-19	Sequence 19, App1
C 42	6.2	62.0	15 10 US-09-880-313A-27	Sequence 27, App1
C 43	6.2	62.0	15 10 US-09-880-313A-29	Sequence 29, App1
C 44	6.2	62.0	15 10 US-09-880-313A-49	Sequence 49, App1
C 45	6.2	62.0	15 10 US-09-880-313A-51	Sequence 51, App1

#### ALIGNMENTS

RESULT 1  
US-10-033-145-59/c  
Sequence 59, Application US/10033145  
Publication No. US2002015151A1  
GENERAL INFORMATION:  
APPLICANT: GENZYME CORPORATION  
APPLICANT: ROBERTS, BRUCE  
APPLICANT: SHANKARA, SRINIVAS  
TITLE OF INVENTION: PREPARATION AND USE OF SUPERIOR VACCINES  
FILE REFERENCE: GA0201C  
CURRENT APPLICATION NUMBER: US/10/033.145  
CURRENT FILING DATE: 2001-11-05  
PRIOR APPLICATION NUMBER: PCT/US99/13800  
PRIOR FILING DATE: 1999-06-18  
NUMBER OF SEQ ID NOS: 2137  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 59  
LENGTH: 10  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-033-145-59

Query Match Best Local Similarity 62.0%; Score 6.2; DB 13; Length 10;  
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 UGCDGSHMND 10  
Db 10 TGCTGGCTAG 1

RESULT 2  
US-10-033-145-312/c  
Sequence 312, Application US/10033145  
Publication No. US2002015151A1  
GENERAL INFORMATION:

```
APPLICANT: GENZYME CORPORATION
APPLICANT: ROBERTS, BRUCE
APPLICANT: SHANKARA, SRINIVAS
TITLE OF INVENTION: PREPARATION AND USE OF SUPERIOR VACCINES
FILE REFERENCE: GAO201C
CURRENT APPLICATION NUMBER: US/10/033,145
CURRENT FILING DATE: 2001-11-05
PRIOR APPLICATION NUMBER: PCT/US99/13800
PRIOR FILING DATE: 1999-06-18
NUMBER OF SEQ ID NOS: 2137
SOFTWARE: Patentin version 3.0
SEQ ID NO 312
LENGTH: 10
TYPE: DNA
ORGANISM: Homo sapiens
US-10-033-145-312
```

```
Query Match      62.0%; Score 6.2; DB 13; Length 10;
Best Local Similarity 40.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 UGCDCGHMND 10
DB      10 TGCAGGTCAA 1
```

```
RESULT 3
US-10-084-700-17
Sequence 17, Application US/10084700
Publication No. US2002016043A1
GENERAL INFORMATION:
APPLICANT: Sealey, Todd
TITLE OF INVENTION: HUBB3 GENE INVOLVED IN HUMAN CANCERS
FILE REFERENCE: PP-01406,004/200130,438D1
CURRENT APPLICATION NUMBER: US/10/084,700
NUMBER OF SEQ ID NOS: 32
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 17
LENGTH: 10
TYPE: DNA
ORGANISM: Homo sapien
US-10-084-700-17
```

```
Query Match      62.0%; Score 6.2; DB 13; Length 10;
Best Local Similarity 40.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 UGCDCGHMND 10
DB      1 TGCAGGTCAA 10
```

```
RESULT 4
US-10-212-476-71
Sequence 71, Application US/10212476
Publication No. US20030157509A1
GENERAL INFORMATION:
APPLICANT: MIRZABEKOV, ANDREI
APPLICANT: GUSCHIN, DMITRY Y.
APPLICANT: SHIK, VALENTINE
APPLICANT: DROBYSHYEV, ALEKSEI
APPLICANT: FOTIN, ALEXANDER
APPLICANT: YERSHOV, GENNADIY
APPLICANT: LYSOV, YU
TITLE OF INVENTION: CUSTOMIZED OLIGONUCLEOTIDE MICROCHIPS THAT CONVERT
TITLE OF INVENTION: MULTIPLE GENETIC INFORMATION TO SIMPLE PATTERNS, ARE
FILE REFERENCE: 21416/90184
CURRENT APPLICATION NUMBER: US/10/212,476
CURRENT FILING DATE: 2002-08-08
PRIOR APPLICATION NUMBER: US/09/261,115
PRIOR FILING DATE: 1999-03-03
```

```
NUMBER OF SEQ ID NOS: 78
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 71
LENGTH: 10
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Customized
US-10-212-476-71
```

```
Query Match      62.0%; Score 6.2; DB 15; Length 10;
Best Local Similarity 40.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 UGCDCGHMND 10
DB      1 TGCAGGTCAA 10
```

```
RESULT 5
US-09-005-243-99/c
Sequence 99, Application US/09005243
Patent No. US20020018763A1
GENERAL INFORMATION:
```

```
APPLICANT: Zeebo, Krzyszina M.
APPLICANT: Boeselman, Robert A.
APPLICANT: Suggs, Sidney V.
APPLICANT: Martin, Francis H.
TITLE OF INVENTION: Stem Cell Factor
NUMBER OF SEQUENCES: 104
CORRESPONDENCE ADDRESS:
```

```
ADDRESSEE: Marshall, O'Toole, Garstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60605-6402
```

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/005,243
FILING DATE:
```

```
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/449,653
FILING DATE: 24-MAY-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/982,255
FILING DATE: 25-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/589,701
FILING DATE: 01-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/573,616
FILING DATE: 24-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/537,198
FILING DATE: 11-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/422,383
FILING DATE: 16-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: Crough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 01017/34465
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
```

TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 99:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-09-005-243-99

Query Match 62.0%; Score 6.2; DB 9; Length 11;  
Best Local Similarity 40.0%; Pred. No. 7.7e+05;  
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 1 UGCDGHNMD 10  
Db 11 TCGAGAGAT 2

RESULT 6  
US-09-224-683-99/C  
Sequence 99, Application US/09224683  
Patent No. US20020031491A1  
GENERAL INFORMATION:  
APPLICANT: Zeebo, Kristina M.  
APPLICANT: Boselman, Robert A.  
APPLICANT: Suggs, Sidney V.  
APPLICANT: Martin, Francis H.  
TITLE OF INVENTION: Stem Cell Factor: Composition Claims  
NUMBER OF SEQUENCES: 104  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/224,683  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/005,893  
FILING DATE: 12-JAN-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/449,653  
FILING DATE: 24-MAY-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/982,255  
FILING DATE: 25-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/589,701  
FILING DATE: 01-OCT-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/573,616  
FILING DATE: 24-AUG-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/537,198  
FILING DATE: 11-JUN-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/422,383  
FILING DATE: 16-OCT-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Clough, David W.  
REGISTRATION NUMBER: 36,107  
REFERENCE/DOCKET NUMBER: 01017/35136

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 99:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-09-224-683-99

Query Match 62.0%; Score 6.2; DB 9; Length 11;  
Best Local Similarity 40.0%; Pred. No. 7.7e+05;  
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 1 UGCDGHNMD 10  
Db 11 TCGAGAGAT 2

RESULT 7  
US-09-772-719-75  
Sequence 75, Application US/09772719  
Patent No. US20020137910A1  
GENERAL INFORMATION:  
APPLICANT: Zavada, Jan  
APPLICANT: Pastorekova, Silvia  
APPLICANT: Pastorek, Jaromir  
TITLE OF INVENTION: MN Gene and Protein  
NUMBER OF SEQUENCES: 86  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Leona L. Lauder  
STREET: 369 Pine Street  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/772,719  
FILING DATE: 30-JAN-2001  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/485,049  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/260,190  
FILING DATE: 15-JUN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Lauder, Leona L.  
REGISTRATION NUMBER: 30,863  
REFERENCE/DOCKET NUMBER: D-0021.3E  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-981-2034  
TELEFAX: 415-981-0332  
INFORMATION FOR SEQ ID NO: 75:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
DESCRIPTION: 5' donor consensus splice sequence  
US-09-772-719-75

Query Match 62.0%; Score 6.2; DB 9; Length 11;  
Best Local Similarity 40.0%; Pred. No. 7.7e+05;  
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 UGCDGSHND 10  
:||||: :  
Db 1 TGCTGGTGAG 10

RESULT 8  
US-09-918-715-162/c  
; Sequence 162, Application US/09918715  
; Publication No. US20030017157A1  
; GENERAL INFORMATION:  
; APPLICANT: Bert Vogelstein  
; APPLICANT: Brad St. Croix  
; APPLICANT: Kenneth Kinzler  
; TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS  
; FILE REFERENCE: 1107.00134  
; CURRENT APPLICATION NUMBER: US/09/918,715  
; PRIOR FILING DATE: 2001-08-01  
; PRIOR APPLICATION NUMBER: 60/222,599  
; PRIOR FILING DATE: 2000-08-02  
; PRIOR APPLICATION NUMBER: 60/224,360  
; PRIOR FILING DATE: 2000-08-11  
; PRIOR APPLICATION NUMBER: 60/282,850  
; PRIOR FILING DATE: 2000-04-11  
; NUMBER OF SEQ ID NOS: 358  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 162  
; LENGTH: 11  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-918-715-162

Query Match 62.0%; Score 6.2; DB 10; Length 11;  
Best Local Similarity 40.0%; Pred. No. 7.7e+05;  
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 UGCDGSHND 10  
:||||: :  
Db 10 TGCGAGTTAT 1

RESULT 9  
US-09-967-237-75  
; Sequence 75, Application US/09967237  
; Publication No. US20030049828A1  
; GENERAL INFORMATION:  
; APPLICANT: Zavada, Jan  
; APPLICANT: Pastorekova, Silvia  
; APPLICANT: Pastorek, Jaromir  
; TITLE OF INVENTION: MN Gene and Protein  
; FILE REFERENCE: D-0021.SB-2  
; CURRENT APPLICATION NUMBER: US/09/967,237  
; PRIOR FILING DATE: 2001-09-27  
; PRIOR APPLICATION NUMBER: 09/178,115  
; PRIOR FILING DATE: 1998-10-23  
; NUMBER OF SEQ ID NOS: 116  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 75  
; LENGTH: 11  
; TYPE: DNA  
; ORGANISM: HUMAN  
US-09-967-237-75

Query Match 62.0%; Score 6.2; DB 10; Length 11;  
Best Local Similarity 40.0%; Pred. No. 7.7e+05;  
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 UGCDGSHND 10  
:||||: :  
Db 1 TGCTGGTGAG 10

RESULT 10

US-10-314-322-303/c  
; Sequence 303, Application US/10314322  
; Publication No. US2003022911A1  
; GENERAL INFORMATION:  
; APPLICANT: Heber-Katz, Ellen  
; TITLE OF INVENTION: Compositions and Methods for Wound  
; FILE REFERENCE: 000486.00016  
; CURRENT APPLICATION NUMBER: US/10/314,322  
; PRIOR FILING DATE: 2002-12-09  
; PRIOR APPLICATION NUMBER: US 60/074,737  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/097,937  
; PRIOR FILING DATE: 1998-08-26  
; PRIOR APPLICATION NUMBER: US 60/102,051  
; PRIOR FILING DATE: 1998-09-28  
; PRIOR APPLICATION NUMBER: US 09/249,155  
; PRIOR FILING DATE: 1999-02-12  
; NUMBER OF SEQ ID NOS: 346  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 303  
; LENGTH: 11  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-10-314-322-303

Query Match 62.0%; Score 6.2; DB 15; Length 11;  
Best Local Similarity 40.0%; Pred. No. 7.7e+05;  
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 UGCDGSHND 10  
:||||: :  
Db 11 TGCTGGAGAG 2

RESULT 11  
US-10-203-969A-5/c  
; Sequence 5, Application US/10203969A  
; Publication No. US20040110224A1  
; GENERAL INFORMATION:  
; APPLICANT: Puljk, Mouter C.  
; APPLICANT: Slocostre, Jelle W.  
; TITLE OF INVENTION: Segment synthesis  
; FILE REFERENCE: B50200US00  
; CURRENT APPLICATION NUMBER: US/10/203,969A  
; PRIOR FILING DATE: 2003-07-07  
; PRIOR APPLICATION NUMBER: EP 00200536.1  
; PRIOR FILING DATE: 2000-02-16  
; PRIOR APPLICATION NUMBER: PCT/NL01/00131  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 660  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 11  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: unspecified  
; OTHER INFORMATION: biotinylated DNA  
; NAME/KEY: misc feature  
; LOCATION: (1)-(11)  
US-10-203-969A-5

Query Match 62.0%; Score 6.2; DB 17; Length 11;  
Best Local Similarity 40.0%; Pred. No. 7.7e+05;  
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 UGCDGSHND 10  
:||||: :  
Db 10 TGCGGGTGCG 1

```
RESULT 12
US-10-450-797-357
; Sequence 357, Application US/10450797
; Publication No. US20040142335A1
; GENERAL INFORMATION:
; APPLICANT: Petersohn, Dirk
; APPLICANT: Conradt, Marcus
; APPLICANT: Hofmann, Kay
; TITLE OF INVENTION: METHOD FOR DETERMINING SKIN STRESS OR SKIN AGEING IN VITRO
; FILE REFERENCE: HENK-0041
; CURRENT APPLICATION NUMBER: US/10/450,797
; PRIOR FILING DATE: 2003-12-04
; PRIOR APPLICATION NUMBER: PCT/EP01/15178
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: DE 101 00 121.5
; NUMBER OF SEQ ID NOS: 1435
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 357
; LENGTH: 11
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-450-797-357

Query Match      62.0%; Score 6.2; DB 17; Length 11;
Best Local Similarity 40.0%; Pred. No. 7.7e+05;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY      1 UGCDGHNMD 10
      :||:|:|:|:
Db      2 TGCTGGAGAA 11

RESULT 13
US-10-450-797-421/c
; Sequence 421, Application US/10450797
; Publication No. US20040142335A1
; GENERAL INFORMATION:
; APPLICANT: Petersohn, Dirk
; APPLICANT: Conradt, Marcus
; APPLICANT: Hofmann, Kay
; TITLE OF INVENTION: METHOD FOR DETERMINING SKIN STRESS OR SKIN AGEING IN VITRO
; FILE REFERENCE: HENK-0041
; CURRENT APPLICATION NUMBER: US/10/450,797
; PRIOR FILING DATE: 2003-12-04
; PRIOR APPLICATION NUMBER: PCT/EP01/15178
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: DE 101 00 121.5
; NUMBER OF SEQ ID NOS: 1435
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 421
; LENGTH: 11
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-450-797-421

Query Match      62.0%; Score 6.2; DB 17; Length 11;
Best Local Similarity 40.0%; Pred. No. 7.7e+05;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY      1 UGCDGHNMD 10
      :||:|:|:|:
Db      10 TGCAGGTCAT 1

RESULT 14
US-10-450-797-663/c
; Sequence 663, Application US/10450797
; Publication No. US20040142335A1
; GENERAL INFORMATION:
; APPLICANT: Petersohn, Dirk
```

```
; APPLICANT: Conradt, Marcus
; APPLICANT: Hofmann, Kay
; TITLE OF INVENTION: METHOD FOR DETERMINING SKIN STRESS OR SKIN AGEING IN VITRO
; FILE REFERENCE: HENK-0041
; CURRENT APPLICATION NUMBER: US/10/450,797
; PRIOR FILING DATE: 2003-12-04
; PRIOR APPLICATION NUMBER: PCT/EP01/15178
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: DE 101 00 121.5
; NUMBER OF SEQ ID NOS: 1435
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 663
; LENGTH: 11
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-450-797-663

Query Match      62.0%; Score 6.2; DB 17; Length 11;
Best Local Similarity 40.0%; Pred. No. 7.7e+05;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY      1 UGCDGHNMD 10
      :||:|:|:|:
Db      11 TGCAGGACAG 2

RESULT 15
US-10-175-608-99/c
; Sequence 99, Application US/10175608
; Publication No. US20040181044A1
; GENERAL INFORMATION:
; APPLICANT: Zeebo, Kristina M.
; APPLICANT: Suggs, Sidney V.
; APPLICANT: Bosseman, Robert A.
; TITLE OF INVENTION: Stem Cell Factor
; NUMBER OF SEQUENCES: 104
; CORRESPONDENCE ADDRESS:
; ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/175,608
; FILING DATE: 16-Oct-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/635,249
; FILING DATE: 07-AUG-2000
; APPLICATION NUMBER: 09/486,546
; FILING DATE: 24-MAY-1995
; APPLICATION NUMBER: 08/172,329
; FILING DATE: 21-DEC-1993
; APPLICATION NUMBER: 07/982,255
; FILING DATE: 25-NOV-1992
; APPLICATION NUMBER: 07/684,535
; FILING DATE: 10-APR-1991
; APPLICATION NUMBER: 09/589,701
; FILING DATE: 10-OCT-1991
; APPLICATION NUMBER: 07/573,616
; FILING DATE: 24-AUG-1990
; APPLICATION NUMBER: 07/537,198
; FILING DATE: 11-JUN-1990
; APPLICATION NUMBER: 07/422,383
; FILING DATE: 16-OCT-1989
```

ATTORNEY/AGENT INFORMATION:  
NAME: Clough, David W.  
REGISTRATION NUMBER: 36,107  
REFERENCE/DOCKET NUMBER: 01017/35199  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 99:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 99:  
US-10-175-608-99

Query Match 62.0%; Score 6.2; DB 17; Length 11;  
Best Local Similarity 40.0%; Pred. No. 7.7e+05;  
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;  
QY 1 UGCDGGHMD 10  
:||:|:|:|:  
Db 11 TCCAGGAGAT 2

Search completed: October 19, 2004, 16:13:07  
Job time: 261 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 19, 2004, 14:44:01 ; Search time 2180 Seconds  
(without alignments)  
167.155 Million cell updates/sec

Title: US-09-762-648-4  
Perfect score: 10  
Sequence: 1 UCGCGHMD 10

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues  
Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gsa1:\*  
9: gb\_gsa2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6.2	62.0	10	CL437053	CL437053 PST4390-N
2	6.2	62.0	13	AJ594173	AJ594173 Arabidops
3	6.2	62.0	19	AZ422762	AZ422762 IM0201P12
4	6.2	62.0	19	AZ791433	AZ791433 2M0041A24
5	6.2	62.0	19	AZ800056	AZ800056 2M0057E20
6	6.2	62.0	19	CL435787	CL435787 PST1409-N
7	6.2	62.0	20	AZ308311	AZ308311 IM0011J12
8	6.2	62.0	20	AZ590589	AZ590589 IM0400G15
9	6.2	62.0	20	AZ630221	AZ630221 IM0483K12
10	6.2	62.0	20	AZ789903	AZ789903 2M0038F15
11	6.2	62.0	21	CO790328	CO790328 NT0009B_E1
12	6.2	62.0	21	AZ594960	AZ594960 IM04007E6
13	6.2	62.0	21	AZ776814	AZ776814 2M0010I17
14	6.2	62.0	21	AZ819369	AZ819369 2M0089O21
15	6.2	62.0	22	AZ307716	AZ307716 IM0009N02
16	6.2	62.0	22	AZ309912	AZ309912 IM0017P15
17	6.2	62.0	22	AZ416988	AZ416988 IM0192P23
18	6.2	62.0	22	AZ501345	AZ501345 IM0340I11
19	6.2	62.0	22	AZ610074	AZ610074 IM0345J11
20	6.2	62.0	22	AZ822589	AZ822589 2M0095P21
21	6.2	62.0	23	BG927951	BG927951 HNC45-1-F
22	6.2	62.0	23	AZ472899	AZ472899 IM0288L22
23	6.2	62.0	23	AZ632990	AZ632990 IM0487M19
24	6.2	62.0	23	AZ810074	AZ810074 2M0074J19

c	25	6.2	62.0	23	8	AZ977634	AZ977634 2M0253C10
c	26	6.2	62.0	23	9	TA121C07P	TA121C07P T. brucei
c	27	6.2	62.0	23	9	TA278H12P	TA278H12P T. brucei
c	28	6.2	62.0	23	9	CL670300	CL670300 PRI01610
c	29	6.2	62.0	23	9	AG195455	AG195455 Pan trogl
c	30	6.2	62.0	24	1	AL045569	AL045569 DKF2p434M
c	31	6.2	62.0	24	5	BO589506	BO589506 E01261-1-0
c	32	6.2	62.0	24	8	AZ308017	AZ308017 IM0010M05
c	33	6.2	62.0	24	8	AZ492799	AZ492799 IM0327B10
c	34	6.2	62.0	24	8	AZ505865	AZ505865 IM0346C18
c	35	6.2	62.0	24	8	AZ821309	AZ821309 2M0094G05
c	36	6.2	62.0	24	9	TA120A03P	TA120A03P T. brucei
c	37	6.2	62.0	24	9	AG187929	AG187929 Pan trogl
c	38	6.2	62.0	25	1	AA880566	AA880566 vx41C05_r
c	39	6.2	62.0	25	1	AA923337	AA923337 o14e07.8
c	40	6.2	62.0	25	1	AA974358	AA974358 oq14E08.8
c	41	6.2	62.0	25	5	BO594927	BO594927 S015257-0
c	42	6.2	62.0	25	8	AZ414090	AZ414090 IM0188F21
c	43	6.2	62.0	25	8	AZ510410	AZ510410 IM0354H21
c	44	6.2	62.0	25	8	AZ769533	AZ769533 IM0570D22
c	45	6.2	62.0	25	9	AJ592851	AJ592851 Arabidops

## ALIGNMENTS

RESULT 1  
CL437053  
LOCUS  
DEFINITION  
PST4390-NR.Seg MICB1 Mus musculus genomic clone PST4390-NR.Seg,  
genomic survey sequence.  
CL437053  
VERSION  
CL437053.1 GI:45572462  
KEYWORDS  
SOURCE  
ORGANISM  
Mus musculus (house mouse)  
GSS.

## REFERENCE

AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Mammalian Functional Genomics Centre  
Manitoba Institute of Cell Biology, University of Manitoba  
ON5029, 675 McDermot Ave, Winnipeg, MB R3E 0V9, Canada  
Tel: 204 787 2133  
Fax: 204 787 2190  
Email: hicksg@cc.umanitoba.ca

## FEATURES

Class: Gene Trap.  
Location/Qualifiers  
1..10  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="129 sv"  
/db\_xref="taxon:10090"  
/clone="PST4390-NR.Seg"  
/sex="Male"  
/cell\_type="Embryonic stem cell"  
/cell\_line="DJ3 (J1 subclone)"  
/clone\_lib="MICB1"  
/note="Vector: U3NeosVI"

## ORIGIN

Query Match 62.0%; Score 6.2; DB 9; Length 10;  
Best Local Similarity 40.0%; Pred. No. 3.3e+06;  
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 UGCDDGHNMD 10  
 Db 10 TGCAGGACAT 1

RESULT 2  
 LOCUS AJS94173  
 DEFINITION Arabidopsis thaliana T-DNA flanking sequence, left border, clone 393H08, genomic survey sequence.

ACCESSION AJS94173  
 VERSION AJS94173.1 GI:37943797  
 KEYWORDS GSS; left border; T-DNA flanking sequence.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1  
 Brunaud, V., Balzerque, S., Dubreucq, B., Aubourg, S., Samson, F., Chauvin, S., Bechold, N., Ctraud, C., Derose, R., Pelletier, G., Lepoint, L., Caboche, M. and Lecharny, A.  
 T-DNA integration into the Arabidopsis genome depends on sequences of pre-insertion sites  
 EMBO Rep. 3 (12), 1152-1157 (2002)

JOURNAL MEDLINE 2263535  
 PUBMED 1246556  
 REFERENCE 2 (bases 1 to 13)  
 Balzerque, S.  
 Direct Submision  
 Submitted (23-OCT-2003) Balzerque S., UMRGV, INRA/CNRS, 2 rue Gaston Cremieux, 91057 Evry cedex, FRANCE  
 PCR was performed on DNA from transformants of Arabidopsis thaliana plants from INRA (Versailles). The DNA fragment(s) resulting from the PCR were directly sequenced from the left or the right border to determine the genomic sequence flanking the insertion. T-DNA derived sequences were removed. Information to order the corresponding mutant line and a link to a database providing a graphical display of the insertion site are available at <http://dbgap.versailles.inra.fr/publiclines/>. This sequence has been generated in the framework of the French plant genomes program 'genoplante' (<http://www.genoplante.com> and <http://genoplante-info.infbioingen.fr>).  
 Location/Qualifiers  
 1..13  
 /organism="Arabidopsis thaliana"  
 /mol\_type="genomic DNA"  
 /cultivar="Wassiliewskija"  
 /db\_xref="taxon:3702"  
 /clone="393H08"  
 /clone\_lib="Arabidopsis thaliana T-DNA insertion lines"  
 1..13  
 /note="T-DNA flanking sequence  
 left border"

ORIGIN  
 Query Match 62.0%; Score 6.2; DB 9; Length 13;  
 Best Local Similarity 40.0%; Pred. No. 3.4e+06;  
 Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 UGCDDGHNMD 10  
 Db 1 TGCCTGGTGA 10

RESULT 3  
 LOCUS AZ422762  
 DEFINITION Ino201P1R Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
 clone UUGC1M0201P12 R, genomic survey sequence.

ACCESSION AZ422762  
 VERSION AZ422762.1 GI:10546871

KEYWORDS GSS.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 19)  
 Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausen, A. and Wright, D. Weis, R.  
 Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts  
 Unpublished (2000)

TITLE Contact: Robert B. Weis  
 University of Utah  
 University of Utah  
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA  
 Tel.: 801 585 5606  
 Fax: 801 585 7177  
 Email: ddunn@genetics.utah.edu  
 Insert Length: 10000 Std Error: 0.00  
 Plate: 0201 row: P column: 12  
 Seq primer: CACACGAAACGCTATGAC  
 Class: plasmid ends  
 High quality sequence stop: 19.  
 Location/Qualifiers  
 1..19  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UUGC1M0201P12"  
 /sex="Male"  
 /lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
 /clone\_lib="Mouse 10kb plasmid UUGC1M library"  
 /note="Vector: PWD42ny; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (GI:4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN  
 Query Match 62.0%; Score 6.2; DB 8; Length 19;  
 Best Local Similarity 40.0%; Pred. No. 3.5e+06;  
 Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 UGCDDGHNMD 10  
 Db 18 TGCAGGCACA 9

RESULT 4  
 LOCUS AZ791433  
 DEFINITION 2M0041A24F Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
 clone UUGC2M0041A24 F, genomic survey sequence.

ACCESSION AZ791433  
 VERSION AZ791433.1 GI:12934312



**KEYWORDS**  
GSS.  
Mus musculus (house mouse)

**SOURCE**  
Mus musculus

**ORGANISM**  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 19)

**REFERENCE**  
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Rellily, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weis, R.  
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts  
Unpublished (2000)

**TITLE**  
Unpublished (2000)

**JOURNAL**  
Contact: Robert B. Weiss  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0041 row: A column: 24  
Seq primer: CCGTGTAAACGACGCGCAGT  
Class: plasmid ends

**FEATURES**  
High quality sequence stop: 19.  
Location/Qualifiers  
1..19  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUCG2M0041A24"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/clone\_lib="Mouse 10kb plasmid UUCG1M library"  
/note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

**ORIGIN**  
Query Match 62.0%; Score 6.2; DB 8; Length 19;  
Best Local Similarity 40.0%; Pred. No. 3.5e+06;  
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;  
QY 1 UGCDCGHMMD 10  
:|:|:|:|:|:|:  
Db 6 TCGAGGCACT 15

**RESULT 5**  
AZ800056 19 bp DNA linear GSS 16-FEB-2001  
LOCUS 2M0057820R Mouse 10kb plasmid UUCG1M library Mus musculus genomic clone UUCG2M0057820 R, genomic survey sequence.  
ACCESSION AZ800056  
VERSION AZ800056.1 GI:12951797

**KEYWORDS**  
GSS.  
Mus musculus (house mouse)

**SOURCE**  
Mus musculus

**ORGANISM**  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 19)

**REFERENCE**  
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Rellily, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weis, R.  
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts  
Unpublished (2000)

**TITLE**  
Unpublished (2000)

**JOURNAL**  
Contact: Robert B. Weiss  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0057 row: B column: 20  
Seq primer: CACACAGGAACAGCTATGACC  
Class: plasmid ends

**FEATURES**  
High quality sequence stop: 19.  
Location/Qualifiers  
1..19  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUCG2M0057E20"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/clone\_lib="Mouse 10kb plasmid UUCG1M library"  
/note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

**ORIGIN**  
Query Match 62.0%; Score 6.2; DB 8; Length 19;  
Best Local Similarity 40.0%; Pred. No. 3.5e+06;  
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;  
QY 1 UGCDCGHMMD 10  
:|:|:|:|:|:|:  
Db 3 TCGTGGAAAT 12

**RESULT 6**  
CL435787 19 bp DNA linear GSS 18-MAR-2004  
LOCUS PST1409-NL-seq MICB1 Mus musculus genomic clone PST1409-NL-seq similar to Spatol, genomic survey sequence.  
ACCESSION CL435787  
VERSION CL435787.1 GI:45569831

KEYWORDS GSS.  
SOURCE Mus musculus (house mouse)  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 19)  
AUTHORS Hicks,G.G.  
TITLE www.escellla.ca  
JOURNAL Unpublished (2002)  
COMMENT Contact: Hicks GG  
Mammalian Functional Genomics Centre  
Manitoba Institute of Cell Biology, University of Manitoba  
ON5029, 675 McDermot Ave, Winnipeg, MB R3E 0V9, Canada  
Tel: 204 787 2133  
Fax: 204 787 2190  
Email: hicks@cc.umanitoba.ca  
U3Neosyl gene trap. Tag generated by plasmid rescue. Additional sequence information and target gene cloning can be generated. ES cell line harboring insertion mutation of target gene is available. Sequence analysis available from  
http://140.193.242.7/escdb/public\_search\_frame.php?PST=PT1409-NL.se

FEATURES  
source  
1..19  
Location/Qualifiers  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="129 sv"  
/db\_xref="taxon:10090"  
/clone="PT1409-NL.seq"  
/sex="Male"  
/cell\_type="Embryonic stem cell"  
/cell\_line="D3H (71 subclone)"  
/clone\_1ib="MICH1"  
/note="Vector: U3Neosyl"

ORIGIN  
Query Match 62.0%; Score 6.2; DB 9; Length 19;  
Best Local Similarity 40.0%; Pred. No. 3.5e+06;  
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;  
QY 1 UGCDGSHMD 10  
4 TGCTGGTAT 13

RESULT 7  
AZ308311 20 bp DNA linear GSS 29-SEP-2000  
LOCUS IM001J12F Mouse 10kb plasmid UGCM library Mus musculus genomic  
DEFINITION clone UGCM001J12 F, genomic survey sequence.  
ACCESSION AZ308311  
VERSION AZ308311.1 GI:10348177  
KEYWORDS GSS.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 20)  
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.  
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts  
JOURNAL Unpublished (2000)  
COMMENT Contact: Robert B. Weiss  
University of Utah Genome Center  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177

Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0011 row: 3 column: 12  
Seq primer: CGTGTAAACGACGGCCAGT  
Class: plasmid ends  
High quality sequence stop: 20.  
Location/Qualifiers  
1..20  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UGCM001J12"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/clone\_1ib="Mouse 10kb plasmid UGCM library"  
/note="Vector: PWD42nv. Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (g1473214|9b|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

FEATURES  
source  
1..20  
Location/Qualifiers  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UGCM001J12"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/clone\_1ib="Mouse 10kb plasmid UGCM library"  
/note="Vector: PWD42nv. Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (g1473214|9b|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN  
Query Match 62.0%; Score 6.2; DB 8; Length 20;  
Best Local Similarity 40.0%; Pred. No. 3.5e+06;  
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;  
QY 1 UGCDGSHMD 10  
11 TGCAGGCCAG 20

RESULT 8  
AZ590589 20 bp DNA linear GSS 13-DEC-2000  
LOCUS IM0400G15F Mouse 10kb plasmid UGCM library Mus musculus genomic  
DEFINITION clone UGCM0400G15 F, genomic survey sequence.  
ACCESSION AZ590589  
VERSION AZ590589.1 GI:11712779  
KEYWORDS GSS.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 20)  
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.  
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts  
JOURNAL Unpublished (2000)  
COMMENT Contact: Robert B. Weiss  
University of Utah Genome Center  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177

Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0400 row: G column: 15  
Seq primer: CCGTGTAAACACGACCGCCACT  
Class: plasmid ends  
High quality sequence stop: 20.  
Location/Qualifiers

## FEATURES

source

1. 20  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUCG1M0483K12"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/clone\_lib="Mouse 10kb plasmid UUCG1M library"  
/note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g1|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

## ORIGIN

Query Match 62.0%; Score 6.2; DB 8; Length 20;  
Best Local Similarity 40.0%; Pred. No. 3.5e+06;

Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 UGCDGHNMD 10  
:|:|:|:|:|:  
Db 20 TGCAGATCA 11

RESULT 9  
A2630221 20 bp DNA linear GSS 13-DEC-2000  
LOCUS 1M0483K12 Mouse 10kb plasmid UUCG1M library Mus musculus genomic  
DEFINITION clone UUCG1M0483K12 R, genomic survey sequence.  
ACCESSION A2630221  
VERSION A2630221.1 GI:11752411  
KEYWORDS GSS.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 20)  
Dunn, D., Aoyagi, A., Barber, M., Baecorn, T., Duval, B., Hamill, C.,  
Irlam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,  
Rellily, M., Rose, M., Rose, R., Stokes, R., Tinney, A., von  
Niederhausern, A. and Wright, D., Weiss, R.  
Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts  
Unpublished (2000)  
Contact: Robert B. Weiss  
University of Utah Genome Center  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177

JOURNAL  
COMMENT

TITLE

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177

Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0483 row: K column: 12  
Seq primer: CACACGAGAAACAGCTATGACC  
Class: plasmid ends  
High quality sequence stop: 20.  
Location/Qualifiers

## FEATURES

source

1. 20  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUCG1M0483K12"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/clone\_lib="Mouse 10kb plasmid UUCG1M library"  
/note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g1|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

## ORIGIN

Query Match 62.0%; Score 6.2; DB 8; Length 20;  
Best Local Similarity 40.0%; Pred. No. 3.5e+06;

Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 UGCDGHNMD 10  
:|:|:|:|:|:  
Db 2 TGTGTGTCA 11

RESULT 10  
A2789903 20 bp DNA linear GSS 16-FEB-2001  
LOCUS 2M0038F15F Mouse 10kb plasmid UUCG1M library Mus musculus genomic  
DEFINITION clone UUCG2M0038F15 F, genomic survey sequence.  
ACCESSION A2789903  
VERSION A2789903.1 GI:12931404  
KEYWORDS GSS.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 20)  
Dunn, D., Aoyagi, A., Barber, M., Baecorn, T., Duval, B., Hamill, C.,  
Irlam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,  
Rellily, M., Rose, M., Rose, R., Stokes, R., Tinney, A., von  
Niederhausern, A. and Wright, D., Weiss, R.  
Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts  
Unpublished (2000)  
Contact: Robert B. Weiss  
University of Utah Genome Center  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177

JOURNAL  
COMMENT

TITLE

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177

Email: ddunn@genetics.utah.edu  
 Insert Length: 10000 Std Error: 0.00  
 Plate: 0038 row: F column: 15  
 Seq primer: CGTGTAAACGACGCCAGT  
 Class: plasmid ends  
 High quality sequence stop: 20.

## FEATURES

## source

1. 20  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UGC2M0038F15"  
 /sex="Male"  
 /lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
 /clone\_lib="Mouse 10kb plasmid UGCGM library"  
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g11473114[gblAF129072.1], a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

## ORIGIN

Query Match 62.0%; Score 6.2; DB 8; Length 20;  
 Best Local Similarity 40.0%; Pred. No. 3.5e+06;  
 Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 UGCDGGHND 10  
 :||:|:|:|:|:  
 Db 8 TCGCAGGCACT 17

RESULT 11  
 C0790328 21 bp mRNA linear EST 05-AUG-2004  
 LOCUS NT0098\_E10 St18-22 Neural tube (NT) Ambystoma mexicanum cDNA 5'  
 DEFINITION similar to hypothetical protein, mRNA sequence.  
 ACCESSION C0790328  
 VERSION C0790328.1 GI:51006299  
 KEYWORDS EST.

SOURCE Ambystoma mexicanum (axolotl)

ORGANISM Ambystoma mexicanum (axolotl)  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Caudata; Salamandroidea; Ambystomidae;  
 Ambystoma.

REFERENCE 1 (bases 1 to 21)  
 Habermann, B., Beblin, A.G., Herklarz, S., Volkmer, M., Eckelt, K.,  
 Pehlke, K., Epperlein, H.H., Schackert, H.K., Wiebe, G. and Tanaka, E.M.  
 An Ambystoma mexicanum EST sequencing project: Analysis of 17,352  
 expressed sequence tags from embryonic and regenerating blastema  
 cDNA libraries  
 Genome Biol. (2004) In press

JOURNAL COMMENT Tanaka Lab  
 Max Planck Institute of Molecular Cell Biology and Genetics,  
 Dresden  
 Pflotenhauserstrasse 108, 01307 Dresden, Germany  
 Tel: 0049 351 210 2620  
 Fax: 0049 351 210 1489

Email: tanaka@mpi-cbg.de  
 Plate: NT0098 row: 10 column: E  
 Seq primer: GCA CAT TAG GCC TAT TTA GGT GAC A.  
 Location/Qualifiers

## FEATURES

## source

1. 21  
 /organism="Ambystoma mexicanum"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:8296"  
 /tissue\_type="Neural Tube, Notochord, Somites"  
 /cell\_type="Includes Neural tube, notochord, somites"  
 /dev\_stage="Stage 18-22"  
 /clone\_lib="St18-22 Neural tube (NT)"  
 /note="Vector: PCWMSport6; Site 1: NotI; Site 2: SalI;  
 Unnormalized cDNA plasmid library prepared by Invitrogen.  
 Size fractionated mRNA plasmid library primed and cloned into  
 NotI-SalI site of PCWMSport6. Bacterial host is  
 EMDH10B-TONA. Average insert size is 1.5 kb.  
 TAG\_LIB=NT"

## ORIGIN

Query Match 62.0%; Score 6.2; DB 7; Length 21;  
 Best Local Similarity 40.0%; Pred. No. 3.5e+06;  
 Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 UGCDGGHND 10  
 :||:|:|:|:|:  
 Db 15 TCGCAGGTGAA 6

RESULT 12  
 A2594960 21 bp DNA linear GSS 13-DEC-2000  
 LOCUS 1M0407E16 Mouse 10kb plasmid UGCGM library Mus musculus genomic  
 DEFINITION clone UGCGM0407E16 F, genomic survey sequence.  
 ACCESSION A2594960  
 VERSION A2594960.1 GI:11717150  
 KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 21)  
 Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,  
 Irlam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,  
 Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von  
 Niederhausen, A. and Wright, D., Weis, R.  
 Mouse whole genome scaffolding with paired end reads from 10kb  
 plasmid inserts  
 Unpublished (2000)

JOURNAL COMMENT Contact: Robert B. Weiss  
 University of Utah Genome Center  
 University of Utah  
 Km. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
 84112, USA  
 Tel: 801 585 5606  
 Fax: 801 585 7177

FEATURES  
 source  
 1. 21  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UGCGM0407E16"  
 /sex="Male"  
 /lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
 /clone\_lib="Mouse 10kb plasmid UGCGM library"  
 /note="Vector: PWD42nv; Purified genomic DNA from M.

Musculus C3H/26J (male) was obtained from the Jackson Laboratory. Mouse DNA Resource (<http://www.jax.org/resource/documents/dnares/>). The DNA was hydrolytically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-kinased with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMDA2 [G14723114]p[AR12072.1], a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

**ORIGIN**

Query Match	62.0%	Score 6.2	DB 8	Length 21
Best Local Similarity	40.0%	Pred. No. 3.5e+06		
Matches	4	Conservative	5	Mismatches 1; Indels 0; Gaps 0;

```
Qy      1 UGCDGGHNM 10
          :||:|: ::
Db      10 TGCTGTACT 19
```

## RESULT 14

LOCUS	AZ819369	21 bp	DNA	linear	GSS 20-FEB-2001
DEFINITION	2M00890201 Mouse 10kb plasmid U06C1M library Mus musculus genomic clone U06C2M0089021 R, genomic survey sequence.				

VERSION	AZ819369.1	GI:12989277
KEYWORDS	GSS.	
SOURCE	Mus musculus (house mouse)	

## ORGANISM

REFERENCE  
1 (bases 1 to 21)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

## AUTHORS

**TITLE** Telam, H., Longacre, S., Mahmud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weiss, R.  
Mouse whole genome scaffolding with paired end reads from 10kx

—

**JOURNAL COMMENT**  
Unpublished (2000)  
Contact: Robert B. Weiss  
University of Utah Genome Center  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA

```
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0089 row: 0 column: 21
Seq primer: CACACGAGAACACGTATGACC
Class: plasmid ends
High quality sequence stop: 21.
```

## FEATURES

```

source
1..21
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUC2M0089021"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F'
/clone_lib="Mouse 10kb plasmid UUC1M library"
/notes="Vector: PWD42nv; Purified genomic DNA from M

```

musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (G14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

## ORIGIN

Query Match 62.0%; Score 6.2; DB 8; Length 21;  
Best Local Similarity 40.0%; Pred. No. 3.5e+06;  
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 UGCDGSHNMD 10  
:|||||:  
Db 4 TGCTGAGACA 13

## RESULT 15

LOCUS AZ307716 22 bp DNA linear GSS 29-SEP-2000  
DEFINITION IM0009N02R Mouse 10kb plasmid UGCGIM library Mus musculus genomic  
clone UGCGIM0009N02 R, genomic survey sequence.  
ACCESSION AZ307716  
VERSION AZ307716.1 GI:10346985  
KEYWORDS GSS.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus

REFERENCE  
AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,  
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,  
Reilly, M., Rose, R., Rose, R., Stokes, R., Tinney, A., von  
Niederhausen, A. and Wright, D., Weis, R.  
Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts  
Unpublished (2000)  
Contact: Robert B. Weiss  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: dunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0009 row: N column: 02  
Seq primer: CACACAGGAAACAGCTATGACC  
Class: plasmid ends  
High quality sequence stop: 22.  
Location/Qualifiers  
1. .22  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UGCGIM0009N02"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/clone\_lib="Mouse 10kb plasmid UGCGIM library"  
/note="Vector: pMD42nv, Purified genomic DNA from M.

JOURNAL COMMENT  
Unpublished (2000)  
Contact: Robert B. Weiss  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: dunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0009 row: N column: 02  
Seq primer: CACACAGGAAACAGCTATGACC  
Class: plasmid ends  
High quality sequence stop: 22.  
Location/Qualifiers  
1. .22  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UGCGIM0009N02"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/clone\_lib="Mouse 10kb plasmid UGCGIM library"  
/note="Vector: pMD42nv, Purified genomic DNA from M.

## ORIGIN

Query Match 62.0%; Score 6.2; DB 8; Length 22;  
Best Local Similarity 40.0%; Pred. No. 3.5e+06;  
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 UGCDGSHNMD 10  
:|||||:  
Db 17 TGCAGAGAG 8

Search completed: October 19, 2004, 16:07:31  
Job time : 2183 secs

musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (G14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."